



UNIVERSIDADE FEDERAL DO RIO GRANDE DO NORTE
CENTRO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA E DE
COMPUTAÇÃO

UFRN
CT
PPGEEC

Análise Estatística e Técnicas de Identificação de Pilotos para Veículos Baja SAE

Tomaz Filgueira Nunes

Orientador: Prof. Dr. Allan de Medeiros Martins

Co-orientador: Prof. Dr. Ivanovitch Medeiros Dantas da Silva

Dissertação de Mestrado apresentada ao
Programa de Pós-Graduação em Engenharia
Elétrica e de Computação da UFRN como
parte dos requisitos para obtenção do título
de Mestre em Ciências.

Número de Ordem do PPgEEC: M540
Natal, RN, Agosto de 2018

Universidade Federal do Rio Grande do Norte - UFRN
Sistema de Bibliotecas - SISBI

Catalogação da publicação na fonte. UFRN - Biblioteca Central Zila Mamede

Nunes, Tomaz Filgueira.

Análise Estatística e Técnicas de Identificação de Pilotos para Veículos Baja SAE / Tomaz Filgueira Nunes. - Natal, 2018.

71f.: il.

Universidade Federal do Rio Grande do Norte. Centro de Tecnologia. Programa de Pós-Graduação em Engenharia Elétrica e de Computação.

Orientador: Prof. Dr. Allan de Medeiros Martins.

Co-orientador: Prof. Dr. Ivanovitch Medeiros da Silva.


1. Análise Estatística de Dados. 2. Baja SAE. 3. Redes Neurais Artificiais. I. Martins, Allan de Medeiros. II. Silva, Ivanovitch Medeiros da. III. Título.

RN/UF/

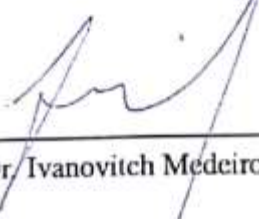
Análise Estatística e Técnicas de Identificação de Pilotos para Veículos Baja SAE

Tomaz Filgueira Nunes

Defesa de Mestrado



Prof. Dr. Allan de Medeiros Martins (orientador) DEE/UFRN



Prof. Dr. Ivanovitch Medeiros da Silva (co-orientador) DCA/UFRN



Prof. Dr. André Felipe Oliveira de Azevedo Dantas UNP



Prof. Dr. Marcelo Henrique Ramalho Nobre IFRN

*“Se for pra desistir, desista
de ser fraco”*

–Autor Desconhecido

*Dedico este trabalho a toda
comunidade Baja*

Agradecimentos

A Deus e a Quequel (*in-memorian*) por terem me guiado nessa longa jornada de dois anos até o sucesso. Necessitei de muita luz espiritual lá de cima para a conclusão deste trabalho.

À minha família por todo apoio e educação que sempre me deram, mesmo longe de casa. Aos amigos próximos: Claudinha, Benilton, Josivan Monte e a Renata de Carvalho pelo companheirismo.

Aos meus amigos de faculdade: Guilherme Bertelli, Ricardo Costa, João Victor Tavares, João Marcos, Márcio Jales, Matheus Torquato, Marie Beltrão, Rafael Teles.

À todos equipe Car-kará de Baja SAE pelo carinho e compreensão mesmo nos momentos difíceis, nas madrugadas, nos testes. Sou muito grato pela realização deste trabalho.

À IAV do Brasil por sempre acreditar na minha capacidade e no meu sucesso.

Resumo

Imergindo-se no contexto dos esportes motorizados, o uso de inteligência artificial se torna uma grande aliada para o bom rendimento de uma equipe de corrida, pois pode extrair características importantes do sistema carro/piloto e assim fornecer *feedbacks* para melhor performance, como já pode ser encontrado em algumas equipes de Formula 1. Partindo deste princípio, este trabalho objetiva a caracterização de pilotos de um veículo *off-road* Baja SAE para que *feedbacks*, durante a temporada de competição, possam ser realizados. Através da parceria com a equipe Car-Kará Baja SAE UFRN, foram selecionados 4 pilotos, com diferentes níveis de experiência em condução de veículo Baja SAE, em 7 pistas de testes. Os dados foram colhidos através de um data logger industrial e analisados de maneira *offline*. A partir da coleta dos dados, fez-se a divisão do vetor de variáveis (RPM, velocidade linear, velocidade angular, aceleração longitudinal e aceleração lateral) em 3 e 5 seções e então foi computada análise estatística (média, desvio padrão, valor máximo e valor mínimo) de cada parte, criando o vetor de características. Esse vetor, por sua vez, foi inserido em uma arquitetura neural artificial de duas camadas ocultas, obtendo uma taxa de classificação, do conjunto de dados reunindo os 4 pilotos, de 97% para o vetor de variáveis com 3 divisões e 93% para o vetor com 5 divisões.

Palavras-chave: Identificação de Sistemas, Análise Estatística de Dados, Baja SAE, Redes Neurais Artificiais.

Abstract

Immersing in the motorsports context, the use of artificial intelligence becomes a great ally to the racing team efficiency by extracting important features from car/driver system and providing feedbacks for a better performance, as it can be found in some Formula 1 teams. From that principle, this work aims to characterize drivers of an off-road Baja SAE vehicle. Through the partnership with the Car-Kará Baja SAE UFRN team, 4 different drivers have been selected in 7 different test tracks. The data has been collected, through an industrial data logger, and analyzed in an offline manner. From the data collection, it has been done a division of the variable vector in 3 and 5 sections and then it has been computed the statistical analysis for each part, creating the feature vectors. That vector was inserted in an artificial neural architecture with two hidden layer, obtaining a classification rate of 97% for the variable vector division in 3 parts and 93% for 5 parts.

Keywords: System Identification, Data Statistical Analysis, Baja SAE, Time Series Models, Artificial Neural Networks.

Sumário

1	Introdução	1
1.1	Aprendizado de Máquina em Esportes a Motor	1
1.2	O Projeto Baja SAE	2
1.3	Objetivo	4
1.4	Organização da Dissertação	5
2	Estado da Arte	7
3	Análise Probabilística e Estatística dos Dados	9
3.1	Introdução a Probabilidade e Variáveis Aleatórias Unidimensionais	9
3.2	Funções de Distribuição Contínuas	10
3.3	Variáveis Aleatórias Multidimensionais	12
3.4	Estimativa de Densidade de Probabilidade	14
3.4.1	Histogramas	14
3.4.2	Kernel	15
4	Redes Neurais Artificiais	19
4.0.1	O Neurônio Artificial	19
4.0.2	Arquitetura de Redes Neurais Artificiais - O Perceptron	21
4.0.2.1	Perceptron de Camada Única	21
4.0.2.2	Perceptron de Múltiplas Camadas	22
4.0.3	Processo de Aprendizado de Redes Neurais	23
5	Metodologia	27
5.1	AIM EV0 4 Data logger	27
5.2	Organização do Procedimento	29
6	Estudo de Caso	31
6.1	Aquisição e organização do Conjunto de Dados	31
6.2	Análise Estatística dos Dados	33

6.2.1	Estimativas das Densidades de Kernel	33
6.2.2	Análise de Correlação dos Dados	34
6.3	Classificação dos Pilotos via Rede Neural	36
7	Resultados	41
8	Conclusão	45
	Referências bibliográficas	47

Lista de Figuras

1.1	Car-Kará UFRN na competição Baja SAE Brasil de 2017.	3
1.2	Exibição das informações ao piloto através de um Display TFT.	3
3.1	Função Densidade de Probabilidade da Distribuição Normal.	12
3.2	A Função de Distribuição Acumulada é determinada pelos limites da área D.	13
3.3	A matriz de correlação é calculada a partir da esperança do produto do vetor de variáveis aleatórias pela sua transposta.	13
3.4	Comparação entre diferentes pontos de início x_o para os histogramas. . . .	15
3.5	Combinação linear dos kernels para a construção da estimativa da distri- buição de densidade.	17
3.6	Estimativa de kernels para diferentes larguras de banda. A figura (a) tem b = 0.2 enquanto que em (b) tem-se largura de banda = 0.8	17
4.1	Modelo do Neurônio Artificial	20
4.2	Neurônios dispostos em uma única camada.	22
4.3	Fronteira de decisão para uma rede perceptron de camada única.	22
4.4	Aquitectura artificial neural disposta em múltiplas camadas.	24
4.5	Processo de Atualização dos Pesos Sinápticos e aprendizado neural. . . .	25
5.1	AIM EVO 4 Data logger.	28
5.2	Fixação do data logger no Baja.	28
6.1	Fluxograma para a montagem do conjunto de dados.	32
6.2	Distribuição Estimada de Kernel dos pilotos nas diversas pistas	35
6.3	Correlação das variáveis para o conjunto de dados utilizando amostras de todos os pilotos nas 7 pistas.	36
6.4	Fluxograma do processo de treinamento da primeira camada oculta da rede neural.	37
6.5	Fluxograma do processo de treinamento das camadas ocultas 1 e 2 da rede neural.	38

7.1	Densidades de probabilidade dos erros de teste	42
-----	--	----

Lista de Quadros e Tabelas

5.1	Variáveis disponíveis no AIM EVO 4	29
5.2	Distribuição dos pilotos nas pistas.	30
6.1	Variáveis Utilizadas	32
6.2	Média e Variância dos treinos dos 4 pilotos em 7 pistas.	33
6.3	Resumo das características da rede neural implementada	39
7.1	Comparativo entre os tempos de treinamento para os algoritmos de Levenberg-Marquardt e Gradient Descent.	42
7.2	Taxa de classificação da Rede Neural para o experimento dos quatro pilotos.	43

Lista de Símbolos e Abreviaturas

$F(x)$	Função de Distribuição Acumulada
$\varepsilon(n)$	Energia do erro de saída da rede neural
μ	média populacional
ρ	Coeficiente de Correlação
σ	Desvio Padrão
σ^2	Variância
$f(x)$	Função Densidade de Probabilidade
w_i	Peso sináptico do neurônio i
x_i	Entrada do neurônio i
ADAS	Advanced Driver Assistance System
CAN	Controller Area Network
COV(X,Y)	Covariância entre as variáveis aleatórias X e Y
DAS	Data Acquisition System
G(.)	Função de Ativação Qualquer
GPS	Global Positioning System
K()	Função Kernel
OBD-II	<i>On Board Diagnostic - versão 2</i>
Rx	Matriz de Correlação
SAE	Society of Automotive Engineers

SSE	Sum of Squared Errors
SVM	Support Vector Machine
TFT LCD	Thin Film Transistor Liquid Crystal Display
UFRN	Universidade Federal do Rio Grande do Norte
V.A	Variável Aleatória

Capítulo 1

Introdução

A evolução da eletrônica automotiva tem chegado a um patamar onde os automóveis tornaram-se sistemas dotados de inteligência. Nesse contexto, muito se fala sobre os carros autônomos tanto é que já se tornou tema de estudo de diversas montadoras, de tal forma que empresas de transporte compartilhado (e.g. Uber) já pensam em ter uma frota de carros sem motoristas (GLANCY, 2015).

O sensoriamento automobilístico foi introduzido na década de 70, com o propósito de atingir níveis satisfatórios para emissões de gases poluentes e se deu ao longo de três “ondas”: a primeira se deu com o crescimento dos sensores aplicados ao sistema de *powertrain*, que consiste no conjunto motor mais transmissão. As “ondas” tiveram foco em reduções de emissões; a segunda foi pela demanda do mercado por performance e segurança; e a terceira foi introduzida com personalização para conforto de sistemas obrigatórios de segurança (GPS, air bags laterais, etc) (FLEMING, 2001).

Diante desse cenário, diversos dispositivos eletrônicos estão embarcados em um automóvel comercial. Segundo Fleming (2008), naquele ano existiam cerca de 100 sensores em um carro de luxo. As unidades de controle eletrônico embarcadas (*ECU - Eletronic Control Unit*) devem ler seus respectivos grupo de sensores e tomar decisões (troca de marcha, anti-travamento de rodas) dependendo de características externas.

1.1 Aprendizado de Máquina em Esportes a Motor

Diversas abordagens podem ser encontradas na literatura e algumas dessas serão comentadas ao longo deste trabalho. Alguns exemplos de aplicações de identificação de sistemas podem ser encontrados nos processos industriais em geral, na economia e inclusive no setor automobilístico. Um tópico pertinente de identificação de sistemas é o aprendizado de máquinas ou *Machine Learning*. Com o uso de técnicas inteligentes como essa, a indústria automotiva está mudando o conceito de automóvel.

ZHANG *et al.* (2010), motivam-se através do exemplo de um usuário com um perfil mais esportivo e outro sendo mais conservador. Dessa forma, o sistema de assistência ao motorista (ADAS) deve ser adaptativo ao grau de dirigibilidade de cada piloto. Além disso, muito se é discutido na literatura sobre a atuação do controlador das transmissões automáticas. Os fabricantes e as grandes montadoras devem sempre fazer um *tradeoff* entre performance e consumo energético. NGO *et al.* (2014) e Zhang 2017, propõem um remapeamento adaptativo do controlador responsável pela troca de marcha, em detrimento da melhora do consumo de combustível. Indo além do nosso dia a dia, também pode-se encontrar *Machine Learning* em competições automobilísticas.

Na principal categoria do automobilismo mundial, a Fórmula 1, o engenheiro de dados deve ficar atento a diversas variáveis. Em um carro de F1, têm-se ao todo cerca de 200 sensores gerando em torno de 300GB de dados (Bill Goldwin, 2017). Dessa forma, a análise de dados e big data estão se tornando assuntos cada vez mais frequentes nessa categoria. Equipes como Mercedes-Benz (Bill Goldwin, 2017) e Renault (Geoff Spencer, 2017) já estão com suas apostas em projetos inteligentes, utilizando aprendizado de máquina, para a temporada de 2018. Com isso, em vez de ter diversos engenheiros olhando para diferentes dados, utiliza-se de inteligência artificial para traçar a melhor estratégia para a equipe, sejam: a entrada no *pitstop*, ritmo de corrida, etc. Dentre as diversas categorias do automobilismo, destaca-se a competição de âmbito acadêmico Baja SAE.

1.2 O Projeto Baja SAE

O Veículo utilizado como material de estudo neste trabalho foi o do tipo Baja SAE da equipe Car-Kará da UFRN, ver Figura 1.1. Ele é um veículo compacto *off-road* desenvolvido com propósito exclusivamente acadêmico. O carro é construído através de estrutura tubular em aço que oferece maior resistência a impactos e que reforça a segurança do condutor em caso de capotamento.

Em relação ao *powertrain*, o motor utilizado no veículo também é padronizado pela SAE: motor de combustão interna, à gasolina, monocilindro de 305 cilindradas produzido pela *Briggs & Stratton* (SAE Brasil, 2018). Já o sistema de transmissão de torque do motor para as rodas é de livre escolha para as equipes. No entanto, é bem comum encontrar transmissões do tipo CVT (*Continuous Variable Transmission*) nos carros tipo Baja SAE, podendo atingir velocidades máximas em torno de 60 km/h.

A eletrônica do veículo é composta por unidades de controle eletrônicos (ECU) dispostas ao longo do carro e são responsáveis pela leitura de diversas variáveis tais como RPM, velocidade linear, nível de combustível, entre outras. As centrais eletrônicas são

Fonte: Autor



Figura 1.1: Car-Kará UFRN na competição Baja SAE Brasil de 2017.

interconectadas através do barramento de comunicação CAN e as informações são exibidas para o piloto através de um display TFT (*Thin Film Transistor*) LCD de 2.8", visto na Figura 1.2.

Fonte: Autor

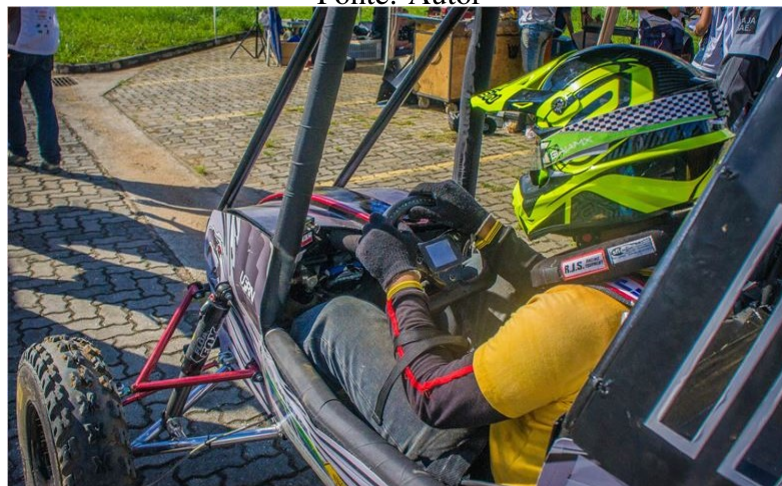


Figura 1.2: Exibição das informações ao piloto através de um Display TFT.

A competição Baja SAE, tanto no escopo nacional quanto internacional, requer que as equipes participantes sigam algumas regras pré-determinadas: o carro não deve exceder 1,90m de altura e deve comportar um piloto com peso máximo de 113,4 kg. Além disso, o carro deve ser robusto para superar condições severas da competição tais como lama, trepidação, chuva, etc. (SAE Brasil, 2018)

Durante a competição as equipes são submetidas a duas etapas de avaliação: estática e dinâmica. A primeira se refere a análise do relatório de projeto, verificação do motor e Inspeções Técnicas e de Segurança do veículo. No entanto, o segundo tipo de avali-

ação trata do desenvolvimento dinâmico do Baja SAE, sendo avaliado em condições de aceleração, velocidade máxima, tração e enduro de resistência.

O projeto Baja SAE na UFRN teve início em 1995 através do departamento de Engenharia Mecânica. A universidade sempre obteve êxito nas competições nacionais de Baja que disputou com suas diversas equipes. No ano de 1996 a equipe Dragões do Sol/UFRN consagrou a melhor dentre as equipes estreantes. No ano seguinte, a mesma equipe conquistou o terceiro lugar, com a UFRN enviando três equipes para a competição.

O auge do projeto Baja SAE na UFRN se deu em 1998 quando a equipe Car-Kará obteve o primeiro lugar na IV competição nacional realizada em Interlagos/SP garantido o direito na participação da etapa mundial realizada nos Estados Unidos e tendo realizado o feito inédito de uma equipe estrangeira ganhar a prova numa competição com outros 86 países.

No ano de 2018, a equipe Car-Kará conseguiu diversos prêmios individuais, obtendo a terceira melhor colocação do Brasil (SAE Brasil, 2018), garantindo a vaga para competição mundial nos Estados Unidos, onde tornou-se a décima melhor equipe do mundo.

Por se tratar de clima de alta competitividade, é necessário a coleta de dados do sistema carro/piloto assim pode-se traçar um *feedback* ao piloto, seja informando sobre comportamentos desnecessários, tais como passagem brusca por obstáculos, curvas imperfeitas, etc; seja por identificação de falhas, através de um software de diagnose embarcada (OBD - *On Board Diagnosis*). Para a realização disso, é necessário um sistemas de aquisição de dados.

Sistemas de aquisição de dados (DAS do inglês) são sistemas que adquirem dados dos veículos para que análises sejam feitas de maneira *off-line*. Eles são bastante utilizados em equipes de competição de esportes motorizados em geral. Tais sistemas gravam e fornecem informações importantes a respeito do comportamento do sistema carro/piloto durante a bateria de testes das equipes (O uso de DAS geralmente são proibidos durante as competições). Os sistemas de aquisições de dados fornecem aos engenheiros de corrida informações suficientemente relevantes para a melhora conjunto carro/piloto para que possam ajudá-los a tomar as melhores decisões.

1.3 Objetivo

Por se tratar de uma equipe de nível internacional, resolve-se aprimorar o sistema de auxílio aos pilotos Car-kará durante as corridas. O caráter competitivo torna os *feedbacks* ao piloto de grande importância, seja informando sobre comportamentos desnecessários ou identificando falhas do veículo. A fim de se analisar a performance do piloto em

situações de competição, é necessário que seja feito de antemão uma identificação do seu perfil de dirigibilidade.

Este trabalho objetiva, em princípio, a caracterização de quatro pilotos de Baja SAE com diferentes níveis de experiência de condução de um veículo Baja SAE (veteranos, intermediário e iniciante), utilizando-se de Redes Neurais Artificiais com duas camadas ocultas. Dessa forma, em um segundo momento, o sistema inteligente pode aprimorar o modo de condução do piloto, fornecendo *feedbacks*, para uma maneira mais competitiva e segura.

1.4 Organização da Dissertação

Esta dissertação está organizada em oito capítulos. No capítulo dois serão expostos alguns trabalhos que utilizam-se da metodologia de classificação de pilotos, relacionando-se, de certa forma, com o que é proposto aqui. Os capítulos três e quatro referem-se a revisão teórica utilizada neste trabalho, de tal forma que no capítulo três encontra-se uma introdução sobre análise estatística dos dados e conceitos de probabilidade, enquanto que no capítulo seguinte estão os conceitos relacionados as Redes Neurais Artificiais.

No capítulo cinco expõe-se a metodologia utilizada para a classificação dos pilotos do veículo Baja SAE. O capítulo seis fala sobre o estudo de caso enquanto que no capítulo sete serão expostos os resultados alcançados. Por fim, o capítulo oito traça algumas conclusões e trabalhos futuros.

Capítulo 2

Estado da Arte

Neste capítulo serão apresentados trabalhos que apresentam grau de semelhança ao proposto neste documento: classificação de motoristas através dos seus comportamentos de direção (*Driving Behavior*).

Propostas de hardware eletrônico também são encontradas na literatura. Em ANDRIA *et al.* (2016), propõe-se o uso de um microcomputador Raspberry pi conectado com um conversor OBD-II/Bluetooth para a análise dos dados de um carro comercial, através da rede CAN (*Controller Area Network*). São utilizados como extrator de características, funções como média e desvio padrão da aceleração longitudinal bem como a razão RPM/Velocidade, para identificar a mudança de marcha. e assim, o perfil do motorista. A saída do classificador é dada como: direção moderada ou direção agressiva.

Além disso, é proposto em NUNES (2016) uma arquitetura de hardware para um veículo do tipo Baja SAE através de controlador Atmel. Foram construídas diversas centrais eletrônicas para a aquisição de dados de sensores tais como tacômetro, velocidade linear, sensor de nível de combustível, entre outros. Os módulos eletrônicos se comunicam entre si através da rede comunicação CAN.

Motivando-se pela identificação de pilotos a partir do comportamento de direção, QUEK; NG (2013) fez um comparativo entre Máquinas de Vetor de Suporte (SVM) e Regressão Logística Multinomial para a classificação de 6 pilotos, de tal forma que os motoristas são identificados um por vez. Utiliza-se como vetor de características a velocidade do carro, o consumo de combustível, aceleração longitudinal e lateral e velocidade angular em curvas. Obteve-se 90% de acurácia com o uso da SVM e 84% para a regressão lógica multinomial.

Mais ainda, encontram-se também trabalhos com a utilização sensores embarcados em *smartphones* como propostos em Van Ly *et al.* (2013) e em EREN *et al.* (2012). No primeiro trabalho, fez-se uso dos sensores inerciais de um *smartphone*, bem como dados reais através da rede CAN de um carro comercial, para a classificação de dois motoristas

a partir da contagem de ocorrência de eventos como frenagem, aceleração e manobra em curvas. Os dados são inseridos em dois modelos: um com aprendizado supervisionado (Máquinas de Vetor de Suporte) e outro não-supervisionado (K-Means). A saída dos modelos é a classificação dos usuários como piloto 1 ou 2. Já na proposta de EREN *et al.* (2012), o classificador de Bayes é utilizado para identificação do comportamento do motorista e tem como vetor de entrada as seguintes variáveis: posição, velocidade, aceleração (positiva e negativa) e o ângulo de deflexão. O vetor de saída é caracterizado pela informação da condição de dirigibilidade, sendo ela "segura" ou "insegura".

Em MESEGUER *et al.* (2013) foi desenvolvida uma plataforma web através de um aplicativo para Android chamado DrivingStyles. O *smartphone* é conectado ao carro através de um scanner OBD-II/*Bluetooth* e envia os dados coletados da rede CAN para a nuvem. Variáveis médias como rpm, velocidade e aceleração são medidas em um período de amostragem de 10 segundos e então analisadas através de duas redes neurais, independentes, que fazem a identificação do tipo de terreno (rodovia de velocidade, estrada, urbano) bem como o grau de agressividade do modo de condução do motorista. A rede neural fornece uma taxa de acerto de 98% para a classificação do tipo do terreno e 77% para o estilo de condução.

Quatro anos depois, em MESEGUER *et al.* (2017), utilizou-se da mesma plataforma web para coletar dados como fluxo de combustível, velocidade linear, entre outras variáveis, para medir o consumo instantâneo de combustível e correlacionando com o estilo de dirigibilidade dos motoristas, mencionando o trabalho anterior.

Este presente trabalho irá abordar conceitos mencionados nesta sessão, como o uso de redes neurais artificiais. Aqui faz-se a separação entre outros trabalhos apresentados na literatura pois é feita a identificação de pilotos de um veículo compacto de competição *off-road* ao invés de terreno urbano.

Capítulo 3

Análise Probabilística e Estatística dos Dados

Neste capítulo serão abordados conceitos introdutórios a probabilidade, como por exemplo variáveis aleatórias (V.A) e funções de distribuição, além da sinergia com a estatística com o estudo por meio das médias e variâncias. Mais ainda, serão abordados pontos referentes a estimativa da distribuição de probabilidade por histogramas e kernel.

3.1 Introdução a Probabilidade e Variáveis Aleatórias Unidimensionais

Existem diversos eventos aleatórios em nosso dia a dia, sejam eles definidos pela análise de ruído de um sinal, sistemas de telecomunicações ou até mesmo o jogar de uma moeda. O conceito de *aleatoriedade* pode ser expressado como algo que não temos informações suficiente, para descrever o comportamento preciso de um sistema. Por exemplo o caso de lançamento de uma moeda, se existisse mais informações sobre a força aplicada sobre ela, orientação inicial, características do material, entre outras, poderíamos prever com exatidão qual seria o valor dela (PISHRO-NIK, 2014).

Seja um experimento aleatório (espaço amostral S) repetido n vezes. Se um evento A desse experimento ocorre n_a vezes, e assumindo que a repetição do experimento é suficientemente grande, pode-se definir $P(A)$ como a probabilidade de ocorrência do evento (A) através da equação 3.1 (PAPOULIS; PILLAI, 2002).

$$P(A) = \lim_{n \rightarrow \infty} \frac{n_a}{n} \quad (3.1)$$

Além disso, alguns axiomas são definidos abaixo:

- $P(A) \geq 0$;

- $P(S) = 1$.

Um dos conceitos mais básicos da análise probabilística são as *Variáveis Aleatórias*. Quando deseja-se descrever o comportamento de um experimento aleatório, geralmente utiliza-se de alguns aspectos como o número de vezes da ocorrência de um determinado evento, por exemplo. Essa "transformação" do comportamento do experimento (espaço amostral S) para um valor numérico referente a cada possibilidade de ocorrência dos eventos (espaço dos números reais) é entendido como variável aleatória, como comentado na Definição 3.1. Uma variável aleatória é retratada por letras maiúsculas (X , Y , Z , etc). Além disso, se um experimento aleatório for discreto (e.g. jogar de uma moeda), a variável aleatória assume valores finitos, caso contrário, quando um experimento for contínuo, assume-se infinitas possibilidades.

Definição 3.1. "Uma variável aleatória X é uma função do espaço amostral para os números reais." (PISHRO-NIK, 2014).

$$X : S \rightarrow \mathbb{R}$$

3.2 Funções de Distribuição Contínuas

O conjunto das probabilidades acumuladas dos possíveis valores de uma variável aleatória contínua (X) é definido pela *Função de Distribuição Acumulada* (F_X), retratada na eq. 3.2. Por outro lado, a *Função Densidade de Probabilidade* (f_X) é a medida de probabilidade por unidade de comprimento Δ . Elas se relacionam entre si por serem funções opostas: uma é a derivada/integral da outra, como mostrado na eq. 3.3 (PISHRO-NIK, 2014).

$$F_X(w) = P(X < w) = \int_{-\infty}^w P_X(u) du \quad (3.2)$$

$$f_X(w) = \frac{P(w < X < w + \Delta)}{\Delta} = \frac{dF_X(w)}{dx}, \text{ Se } F(x) \text{ for diferenciável em } x \quad (3.3)$$

Tendo como base um experimento qualquer e sendo X a variável aleatória desse experimento, pode-se afirmar que o valor esperado (ou esperança) dessa variável aleatória, definido por $E[X]$, será determinado pela eq. 3.4, sendo $f(x)$ uma função de densidade de probabilidade. No entanto, se o experimento contiver um número de amostras/eventos (n) suficientemente grande, o valor esperado $E[X]$ será uma média aritmética simples (μ)

descrita pela eq. 3.5. Além disso, a variância, σ^2 , que é uma medida do grau de dispersão de uma variável aleatória para com a sua média é determinada pela eq. 3.6. A raiz quadrada da variância (σ) é chamada de *desvio padrão* (PAPOULIS; PILLAI, 2002).

$$E[X] = \int_{-\infty}^{\infty} xf(x)dx \quad (3.4)$$

$$E[X] = \mu = \frac{\sum_{i=1}^n x_i}{n} \quad (3.5)$$

$$\sigma^2 = E[X^2] - E[X]^2 = \frac{\sum_{i=1}^n (x_i - \mu)^2}{n - 1} \quad (3.6)$$

Algumas dessas funções de distribuição são usadas com bastante frequência, pois modelam diversos sistemas físicos reais, e, assim, recebem nomes específicos (PISHRO-NIK, 2014): uma dessas distribuições é a *Normal*.

Distribuições normais (ou normais gaussianas) são denotadas matematicamente por $X \sim N(\mu, \sigma^2)$, sendo X a variável aleatória que possui tal distribuição; μ e σ são a média e variância, respectivamente, da distribuição normal dada. Além disso, se $\mu = 0$; $\sigma^2 = 1$ tal distribuição tem nome de *Normal Padronizada* (PAPOULIS; PILLAI, 2002).

No entanto, muitos sistemas físicos que são caracterizados por uma distribuição normal possuem diferentes médias e variâncias. Nesse caso é necessário deslocar essas características para que todas as variáveis aleatórias sejam regidas pela distribuição normal padronizada. Esse processo chama-se *padronização para score* Z e pode ser entendido pela eq. 3.7, sendo $X \sim N(\mu, \sigma^2)$ e $Z \sim N(\mu = 0, \sigma^2 = 1)$ (PAPOULIS; PILLAI, 2002). A densidade de probabilidade da distribuição normal padronizada é regida pela Eq. 3.8 abaixo.

$$Z = X - \frac{\mu}{\sigma} \quad (3.7)$$

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-z^2}{2}\right) \quad (3.8)$$

A Figura 3.1 retrata diferentes distribuições normais. Todas elas possuem média μ iguais a zero, enquanto que a variância dos dados divergem de uma distribuição para outra. Pode-se perceber que quanto maior a variância, mais disperso estão os valores.

Fonte: Autor

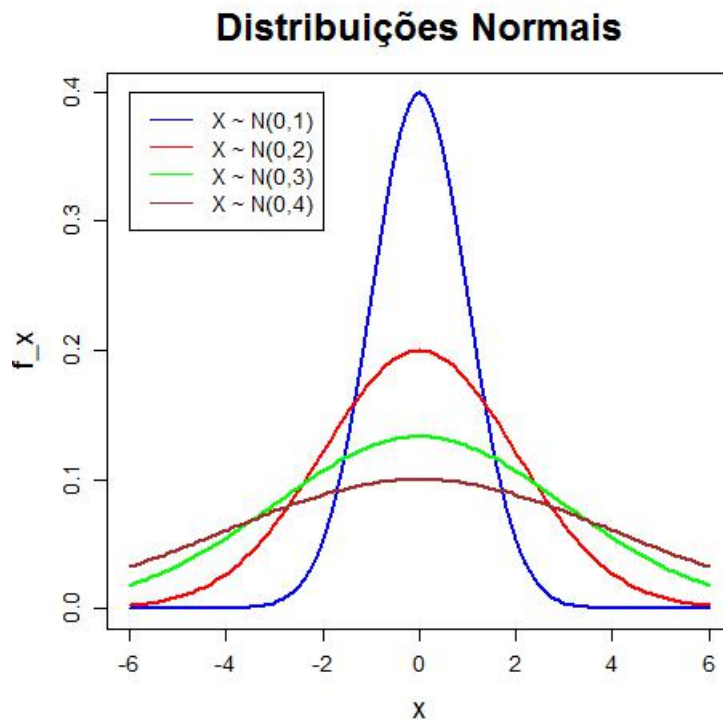


Figura 3.1: Função Densidade de Probabilidade da Distribuição Normal.

3.3 Variáveis Aleatórias Multidimensionais

O conceito de variáveis aleatórias multidimensionais é uma extensão do caso de uma única V.A. Aqui o estudo estatístico é baseado em n funções de distribuição. O caso mais simples é quando têm-se duas V.A. Tomando como um exemplo desse caso, têm-se duas variáveis aleatórias X e Y nas quais deseja-se saber qual é a probabilidade de ocorrência em um determinado ponto ($X=x, Y=y$) de uma região D no plano xy , retratado pela Figura 3.2. A probabilidade está diretamente ligada a Função de Distribuição Acumulada conjunta, como mostrado na eq. 3.9. Esse estudo estatístico é chamado de probabilidade conjunta.

$$F(x,y) = P\{X \leq x, Y \leq y\} \quad (3.9)$$

Indo além no estudo das variáveis aleatórias bidimensionais, existe também o conceito da covariância entre elas, que mede o quão dispersos os dados estão do ponto médio das V.A. (μ_x, μ_y). A covariância é calculada a partir da eq. 3.10 e no caso das variáveis aleatórias serem independentes, ou seja, o comportamento de uma não é influenciado pela

Fonte: (PAPOULIS; PILLAI, 2002)

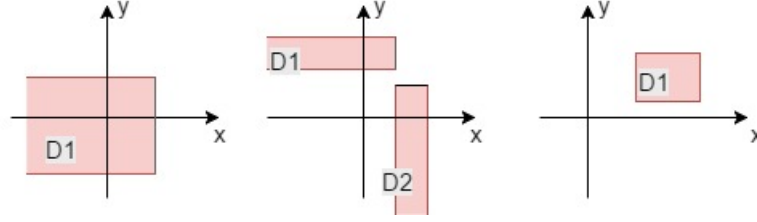


Figura 3.2: A Função de Distribuição Acumulada é determinada pelos limites da área D.

outra, a covariância entre elas é zero (PISHRO-NIK, 2014).

$$COV(X, Y) = E[(X - E[X])(Y - E[Y])] = E[XY] - E[X]E[Y] \quad (3.10)$$

Outro conceito importante é o coeficiente de correlação de variáveis aleatórias bidimensionais que mede o quão relacionadas estão as variáveis aleatórias em relação ao seu comportamento. Em outras palavras pode-se dizer que é quanto o valor de uma variável cresce/diminui de acordo com o comportamento da outra. O coeficiente de correlação é calculado a partir da covariância entre as variáveis (X e Y) e suas respectivas variâncias individuais, como mostrado na eq. 3.11. O coeficiente é normalizado entre o intervalo $[-1, 1]$, sendo o valor de -1 representando a correlação total negativa, enquanto que o valor de 1 representa a correlação total positiva entre as variáveis em questão. O valor de coeficiente igual a zero indica desconexão entre os dados analisados.

$$\rho = \frac{COV(X, Y)}{\sigma_x \sigma_y} \quad (3.11)$$

Seja um vetor com N variáveis aleatórias $X = [X_1, X_2, \dots, X_N]$. Se aplicarmos a operação esperança, sabendo que ela é linear, chegamos ao vetor $E[X] = [EX_1, EX_2, \dots, EX_N]$. A matriz de correlação (R), que contém os coeficientes calculados pela eq. 3.11, pode ser encontrada diretamente através esperança da multiplicação do vetor de variáveis aleatórias pela sua transposta, como mostrado na Figura 3.3.

Fonte: (PISHRO-NIK, 2014)

$$R_x = E[XX^T] = \begin{bmatrix} E[X_1^2] & E[X_1X_2] & \dots & E[X_1X_n] \\ E[X_2X_1] & E[X_2^2] & \dots & E[X_2X_n] \\ \vdots & \vdots & \ddots & \vdots \\ E[X_nX_1] & E[X_nX_2] & \dots & E[X_n^2] \end{bmatrix}$$

Figura 3.3: A matriz de correlação é calculada a partir da esperança do produto do vetor de variáveis aleatórias pela sua transposta.

3.4 Estimativa de Densidade de Probabilidade

Supondo que tenham-se dados coletados de um sistema real que não sabe-se, a priori, qual a distribuição de probabilidade, define-se então o conceito de estimação da densidade como a construção de uma estimativa de distribuição (\hat{f}) que mais se aproxime da curva real (f) e assim algumas conclusões para os dados adquiridos podem ser feitas. A estimativa da distribuição pode ser baseada em duas abordagens.

A primeira abordagem, chamada de paramétrica, pode ser feita através da estimação de parâmetros da função de densidade, por exemplo média μ e variância σ^2 . Por outro lado, quando estima-se a densidade de probabilidade de um sistema físico a partir apenas dos dados obtidos experimentalmente, temos a abordagem não-paramétrica. Neste capítulo, tem-se o foco na segunda abordagem, explicando algumas estimativas como: histograma e Kernel.

3.4.1 Histogramas

Os histogramas são um dos estimadores de densidades mais utilizados, por sua simplicidade. O espaço de dados é dividido em “n” categorias (ou faixa de valores) distintas e que são mostradas visualmente através de barras. Tendo como premissa um ponto inicial em x_0 , uma largura de barra h (fixa para todas as categorias) e sabendo quantos pontos dos dados coletados estão presentes em um determinada categoria x , pode-se estimar a função densidade de probabilidade através da eq. 3.12:

$$\hat{f}(x) = \frac{\text{Números de Amostras } [X_i \in x]}{nh} = \frac{k}{nk} \quad (3.12)$$

No entanto, a simplicidade do uso dos histogramas traz algumas consequências. Em SILVERMAN (1996) menciona-se que um dos problemas em se usar histogramas como estimador de densidade de probabilidade são as descontinuidades entre barras que podem gerar problemas caso sejam necessárias computações de derivadas. Além disso, GRAMACKI (2018) menciona que para a construção de histogramas devemos sempre estar atentos que o ponto inicial (x_0) da primeira barra influencia bastante na forma deste tipo de estimador, modificando assim toda a análise. O mesmo autor conclui que o uso de histogramas é para apenas rápida visualização da distribuição dos casos em um caso de dimensão um ou dois (uni e multivariável).

A Figura 7.1, retrata um pouco do problema da distorção de forma distribuição por histogramas. O conjunto de dados analisados é o acúmulo anual de neve na cidade de Buffalo, nos Estados Unidos. Tendo uma largura de barra fixa em $h=10$ e modificando o

ponto de início x_0 entre os valores $[0, 2, 4, 6, 8]$. O histograma do canto inferior direito retrata a densidade "verdadeira" através da estimativa da densidade de kernel. (GRAMACKI, 2018).

Fonte: (GRAMACKI, 2018)

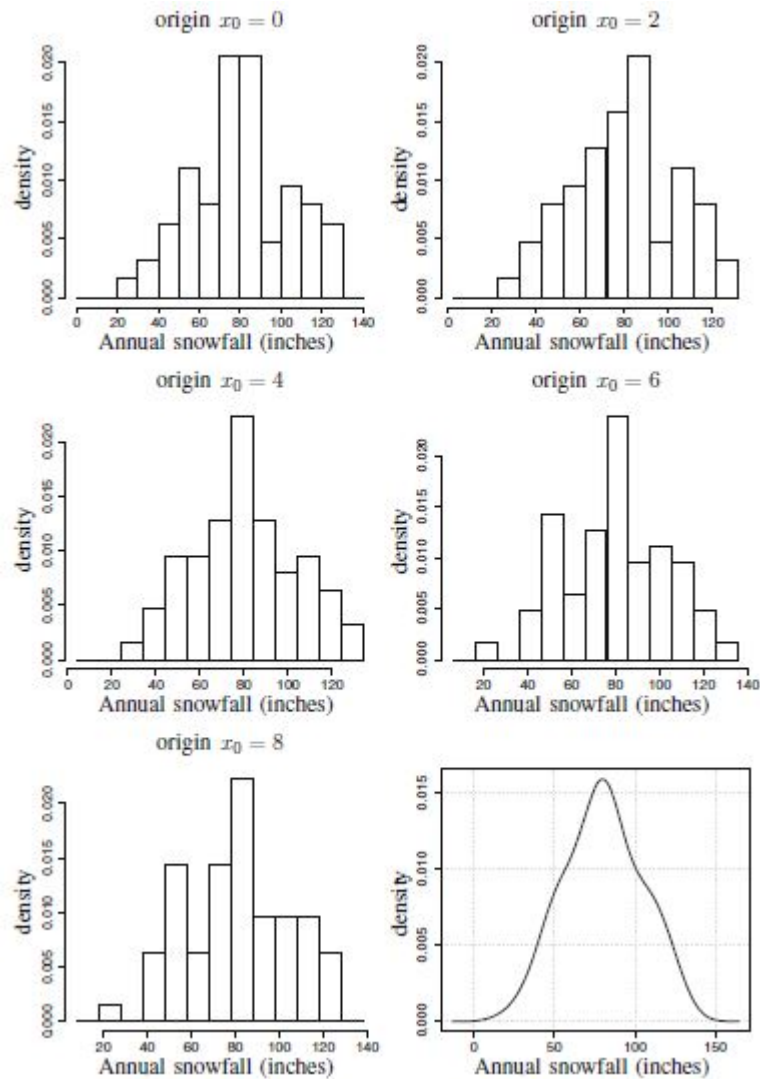


Figura 3.4: Comparação entre diferentes pontos de início x_0 para os histogramas.

3.4.2 Kernel

No contexto da estatística, o kernel é uma função previamente definida que é utilizada, como auxílio, para estimar a distribuição de probabilidades de um conjunto de dados. Partindo da definição de Função Densidade de Probabilidade, uma variável aleatória X

tem densidade f através da eq. 3.13 .

$$f(x) = \lim_{b \rightarrow 0} \frac{P(x-b < X < x+b)}{2b} \quad (3.13)$$

Além disso, o numerador da eq. 3.13 pode ser estimado, para qualquer b , a partir da contagem de elementos que estão no intervalo $[x-b, x+b]$ dividido pelo número total de elementos da amostra, b é comumente chamado de *largura de banda* do kernel. Dessa forma, a equação de densidade de probabilidade estimada é descrita pela eq. 3.14. Reescrevendo-a de uma forma mais transparente, tem-se a eq. 3.15 que explicita o kernel K como uma função (SILVERMAN, 1996).

$$\hat{f}(x) = \frac{\# X_i \in [x-b, x+b]}{2bn} \quad (3.14)$$

$$\hat{f}(x) = \frac{1}{nb} \sum_{i=1}^n K\left(\frac{x-X_i}{b}\right) \quad (3.15)$$

Mais ainda, deve-se ter em mente que as funções de Kernel ($K(\cdot)$ é uma função genérica de kernel) são regidas por algumas propriedades mencionadas abaixo, sendo $u = (x - X_i)$ e k a ordem do kernel. É importante ressaltar que se a segunda suposição é verdadeira, então garante-se que a função $\hat{f}(\cdot)$ também é uma distribuição de probabilidade. Alguns kernels importantes são: Uniforme, Triangular, Gaussiano, entre outros (TURLACH, 1993).

1. $K(\cdot)$ é simétrica. $K[u] = K[-u]$;
2. $\int K(u)du = 1$;
3. $\int u^j K(u)du = 0$, para $j=1, \dots, k-1$
4. $\int u^k K(u)du \neq 0$.

Na prática, a estimativa de distribuição nada mais é do que uma combinação linear de kernels K (que determinam a forma), de tal maneira que em cada ponto analisado é inserida uma distribuição de densidade com largura de banda b (que define o quão largo é a curva). A Figura 3.6 retrata a combinação linear de vários kernels gaussianos, com largura de banda $b = 0,4$, para a construção da estimativa de densidade.

Deve-se fazer um *Trade-off* para a seleção da melhor largura de banda da estimativa. Caso seja selecionado um número elevado, a estimativa não representa fielmente a distribuição original. Este é o caso do que chama-se de *underfitting*. Por outro lado, quando

Fonte: (SILVERMAN, 1996)

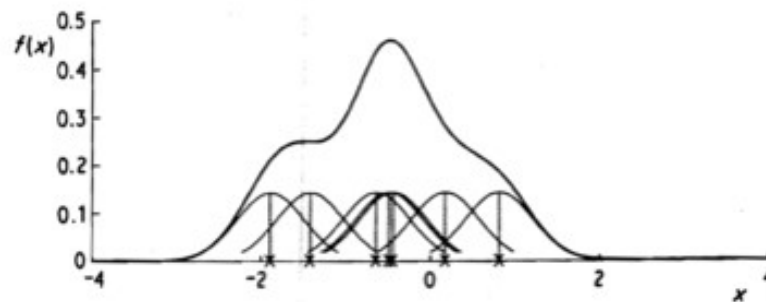


Figura 3.5: Combinação linear dos kernels para a construção da estimativa da distribuição de densidade.

seleciona-se um número b suficientemente pequeno a distribuição começa a ficar bastante variável, representando, em excesso os dados. Aqui chama-se de *overfitting*.

Fonte: (SILVERMAN, 1996)

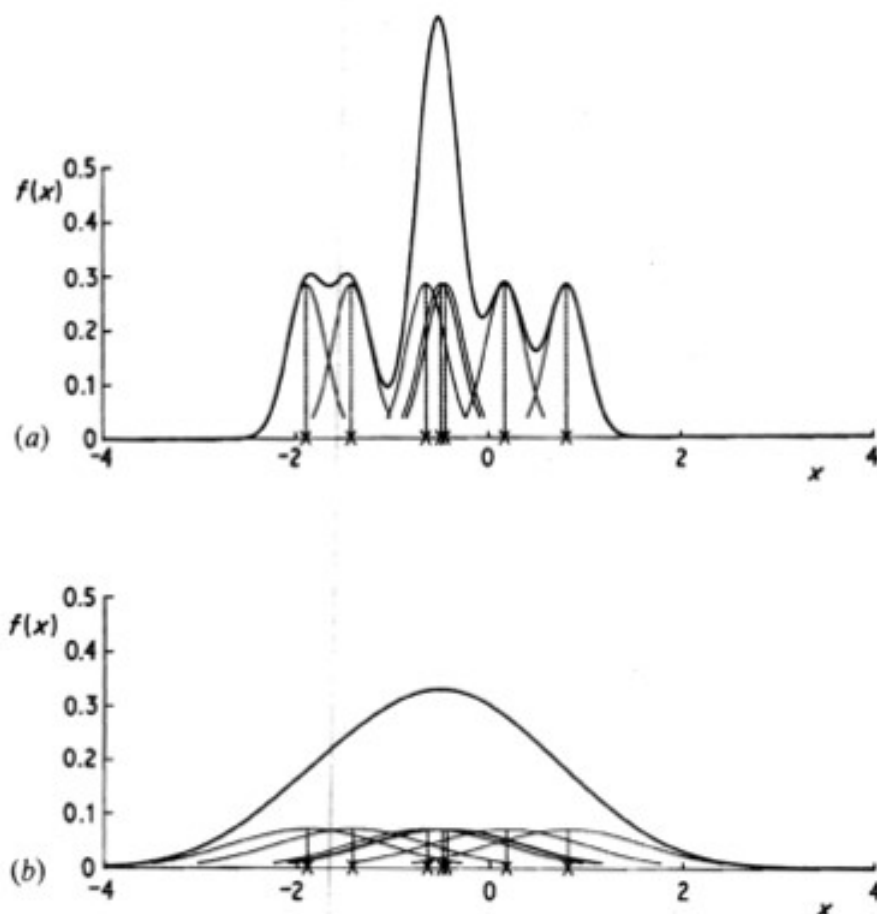


Figura 3.6: Estimativa de kernels para diferentes larguras de banda. A figura (a) tem $b = 0.2$ enquanto que em (b) tem-se largura de banda = 0.8

Capítulo 4

Redes Neurais Artificiais

As Redes Neurais Artificiais são definidas como

"Uma rede neural é um processador maciçamente paralelamente distribuído constituído de unidades de processamento simples, que a propensão natural para armazenar conhecimento experimental e torná-lo disponível para o uso". [...] (HAYKIN, 2001)

Os modelos modernos que conhecemos hoje foram introduzidos na comunidade científica na década de 40 do século passado pelo trabalho pioneiro de McCulloch e Pitts. O primeiro foi um psiquiatra, neuroanatomista e grande curioso sobre a representação do sistema nervoso biológico e o segundo foi um grande nome da matemática que fez parceria com McCulloch, em 1942, com o objetivo de desenvolver um modelo matemático que representasse as conexões do cérebro humano através de estudos da neurofisiologia e lógica.

As Redes Neurais Artificiais têm como objetivo se assemelhar ao cérebro humano em termos de funcionalidades. O conhecimento que a rede adquire com o tempo se dá com o processo de treinamento da mesma e a informação do aprendizado é armazenada nas conexões entre unidades neurais, ou pesos sinápticos. (HAYKIN, 2001).

4.0.1 O Neurônio Artificial

O modelo do neurônio artificial proposto por McCulloch-Pitts é composto por diversos sinais de entradas e uma saída, como pode ser visto na Figura 4.1. Entre as entradas (X_s) e a saída (Y) existem outros elementos básicos que simulam a funcionalidade do neurônio biológico - conjunto de sinapses ou pesos sinápticos (W 's); o somador que faz a operação de soma do produto entre os sinais de entrada e os respectivos pesos sinápticos, ele recebe o nome de campo local induzido (v_k); a função de ativação $G(.)$ que limita o valor do sinal de saída; e o bias que consiste em uma entrada de valor constante, podendo

ser positivo ou negativo, a fim de ajustar a entrada da função de ativação. A relação entre as entradas do neurônio e a saída do operador somador é descrita pela equação. 4.1.

$$v_k = \sum_{i=1}^n x_i w_{ki} + b_k \quad (4.1)$$

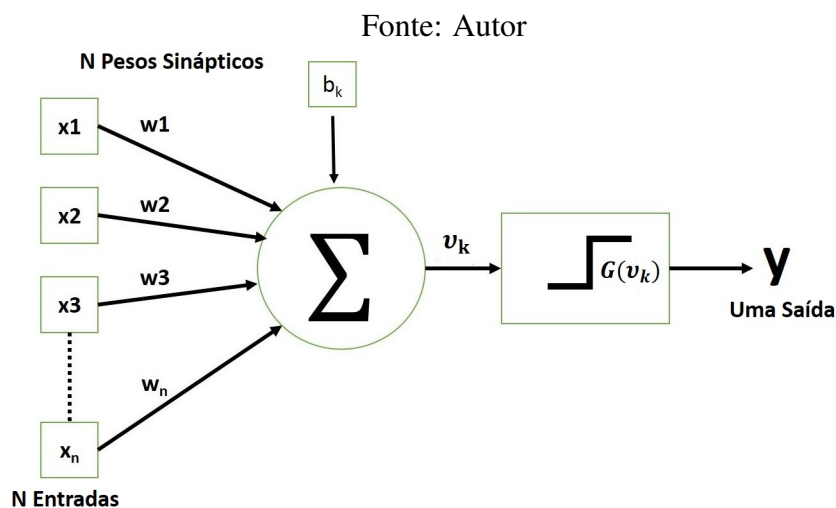


Figura 4.1: Modelo do Neurônio Artificial

A função de ativação por sua vez pode ser definida de diversas formas. No caso específico do modelo artificial proposto por McCulloch-Pitts a saída do neurônio é binária. Isso se dá devido ao uso da função do tipo degrau como função de ativação $G(v_k)$. Outros tipos de funções podem ser como mostrado abaixo (HAYKIN, 2001).

- Linear por partes - Este tipo de função é regido por uma equação de reta e possui saídas proporcionais aos seus sinais de entrada. Sua equação é encontrada em eq. 4.2.
- Sigmoide - A função sigmoide tem formato em 's' e é uma das mais encontradas nas aplicações de redes neurais. A eq. 4.3 define esta função de ativação, onde α é a constante de inclinação da função sigmoide.
- *Softmax* - Este tipo de função de ativação é um caso particular da função sigmoide. Aqui, as saídas da rede neural podem ser interpretadas como probabilidades da classe em questão. O somatório de todas as saídas deve ser igual a 1 (BISHOP, 1995).
- Tangente hiperbólica - Esta é uma extensão da função sinal juntamente com a função sigmoide. Ela gera saídas entre -1 e 1 enquanto que no caso sigmoide apenas entre 0 e 1. Sua equação é definida em eq. 4.4.

$$G(v_k) = \begin{cases} 1, & v_k \geq 0.5 \\ v_k, & +0.5 > v_k > -0.5 \\ 0, & v_k \leq -0.5 \end{cases} \quad (4.2)$$

$$G(v_k) = \frac{1}{1 + \exp(-\alpha v_k)} \quad (4.3)$$

$$G(v_k) = \tanh(v_k) \quad (4.4)$$

O modelo proposto por McCulloch-Pitts possuía pesos sinápticos fixos. Rosenblatt, em 1957, propôs o uso desse modelo com pesos sinápticos ajustáveis por um algoritmo de treinamento. Esse novo modelo passou a ser chamado de *Perceptron de Rosenblatt*.

4.0.2 Arquitetura de Redes Neurais Artificiais - O Perceptron

Como visto nas equações anteriores, um único neurônio gera como sinal de saída valores entre 0 e 1. Dependendo da forma de organização das redes neurológicas artificiais um aumento da performance pode ser identificado, definindo-se uma ou diversas camadas ocultas na arquitetura neural.

4.0.2.1 Perceptron de Camada Única

Em uma distribuição neural de camada única, encontra-se a mais simples arquitetura possível. Neste caso, os neurônios se organizam em coluna de nós que se projeta diretamente para a saída da rede (Y), como visto na Figura 4.2 na qual se vê um exemplo de uma arquitetura simples com quatro neurônios (X_1, X_2, X_3, X_4). É válido ressaltar que a informação se propaga dos nós de entrada (a esquerda) até a camada de saída dos elementos neurais (a direita). O caminho inverso jamais é possível (HAYKIN, 2001).

Tomando como exemplo uma rede perceptron de uma única camada com apenas duas entradas (x_1 e x_2), com base na eq. 4.1 que refere-se ao potencial de ativação do neurônio, chega-se a eq. 4.5, que é a saída do perceptron. A partir dela pode-se perceber que a fronteira de decisão entre as duas classes (A e B) é determinada por uma reta. Conclui-se que o perceptron de uma única camada consegue apenas distinguir padrões que sejam linearmente separáveis por um hiperplano, como mostra a Figura 4.3 (NUNES *et al.*, 2010).

$$Y = \begin{cases} 1, & \text{se } \sum w_i x_i + b \Leftrightarrow w_1 x_1 + w_2 x_2 + b \geq 0; \\ -1, & \text{se } \sum w_i x_i + b \Leftrightarrow w_1 x_1 + w_2 x_2 + b < 0; \end{cases} \quad (4.5)$$

Fonte: Autor

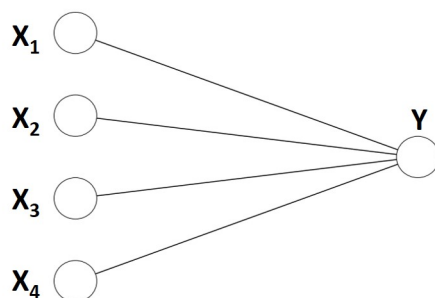


Figura 4.2: Neurônios dispostos em uma única camada.

Fonte: (NUNES *et al.*, 2010)

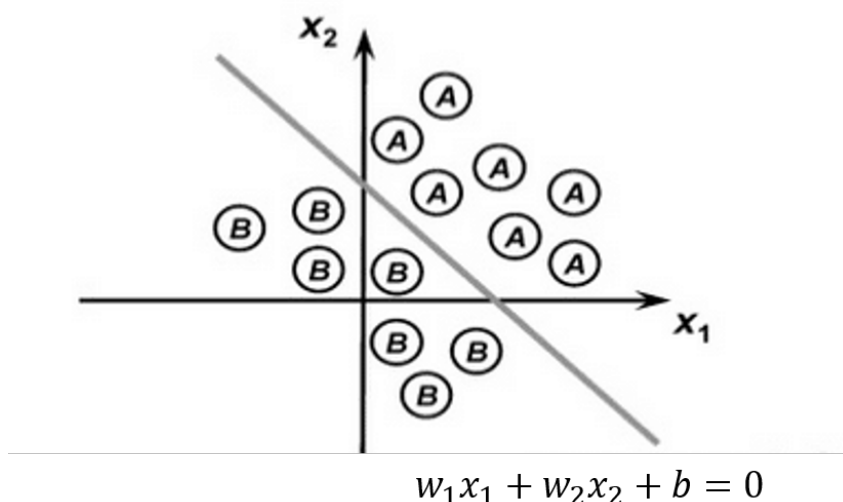


Figura 4.3: Fronteira de decisão para uma rede perceptron de camada única.

Indo além, podem-se encontrar neurônios distribuídos em diversas camadas, aumentando assim o potencial de extração de características do ambiente no qual está inserido a rede. Neste trabalho as camadas ocultas serão denotadas pela letra H.

4.0.2.2 Perceptron de Múltiplas Camadas

Diferenciando da sessão anterior, tem-se a arquitetura de redes neurais em que suas unidades são distribuídas sob várias camadas, como pode ser visto na Figura 4.4. Nela encontram-se duas camadas de neurônios: a camada de entrada contendo os nós de fonte (esta camada não é contabilizada), duas camadas ocultas (H1 e H2) e uma camada de saída com dois neurônios. Neste tipo de rede, podem ser encontradas diversas camadas ocultas que tem como propósito atuar entre os nós de entrada e os de saída eficazmente. Quando tem-se mais de duas camadas ocultas de neurônios, a rede neural é comumente

chamada de Rede Profunda ou *Deep Neural Network* do inglês. O processamento da informação se dá pela seguinte maneira:

1. A informação externa é passada para a rede neural através dos nós de entrada;
2. Os nós de entrada por sua vez estão diretamente ligados a primeira camada oculta da rede neural onde é processado pela soma ponderada do campo induzido e sua saída é determinada pela respectiva função de ativação;
3. A saída da primeira camada oculta está ligada na entrada da segunda camada oculta, se houver, caso contrário a saída da primeira camada oculta passará para a última camada da rede neural.
4. A camada mais à direita do diagrama da Figura 4.4 é a saída da rede. Nela, a informação previamente processada pelas outras camadas chega e, então, é feita sua respectiva função (identificação, classificação, etc)

Com a presença da camada oculta numa rede neural, o sistema torna-se capaz de extrair características de alta complexidade (e.g determinada posição de pixel de uma imagem que contem a forma de uma letra do alfabeto para a sua determinação). No entanto, quanto maior a complexidade do problema, maior será a quantidade de unidades processadoras de informação na camada oculta e, conseqüentemente, o aumento do tempo necessário para treinamento da rede (BISHOP, 1995).

Adicionalmente, Nunes et al. (2010) mencionam que, ao contrário do que ocorre no perceptron de camada única o qual consegue classificar apenas padrões linearmente separáveis, aqui, o classificador resolve o problema de classificação de padrões quando os elementos estejam dentro de uma região convexa.

4.0.3 Processo de Aprendizado de Redes Neurais

Como mencionado anteriormente, o conhecimento que a rede neural adquire com o tempo é armazenado nos pesos sinápticos. Porém, para que a rede chegue em determinado nível de aprendizado, ela precisa passar por algum algoritmo de treinamento. Chollet (2007) menciona que a tarefa de *aprendizagem* refere-se a encontrar os melhores valores para os pesos sinápticos, de todas as camadas, de tal forma que a rede neural consegue mapear cada entrada X com sua respectiva classificação Y. O processo de aprendizagem (atualização dos pesos sinápticos) pode ser feito com ou sem um supervisor.

O supervisor de uma rede neural pode ser interpretado como um bloco que tem conhecimento sobre o ambiente em que a rede está inserida, sendo representado pelos tuplas

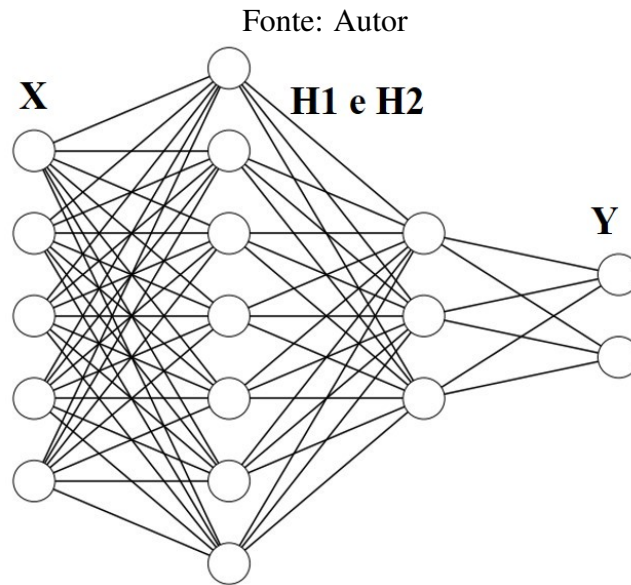


Figura 4.4: Arquitetura artificial neural disposta em múltiplas camadas.

(x, y) de entrada-saída. Inicialmente, os valores dos pesos das interconexões neurais são definidos como números aleatórios e a arquitetura neural implementa apenas transformações no sinal de entrada gerar um valor de classificação na saída. É necessário então medir o quão distante o valor de saída (Y) está do valor desejado (conhecido pelo supervisor) (CHOLLET, 2007).

A fim de minimizar a distância entre esses valores de saída, é implementada a *função custo* da rede neural. Ela tem como base algumas métricas de erro, como por exemplo a soma dos erros quadráticos (SSE - *Sum of Squared Errors*) já descrita pela eq. 4.6. Dessa forma, o supervisor atua a cada iteração com o intuito de minimizar o valor da função custo, atualizando os pesos sinápticos de tal forma que o gradiente do erro se estabilize e seja mínimo. Quando a rede obtiver todo o conhecimento do ambiente, através do supervisor e da maneira mais completa possível, diz que a emulação é ótima no sentido estatístico. A partir de então a rede é deixada livremente, sem o bloco de supervisão, para lidar com o ambiente (HAYKIN, 2001). A Figura 4.5 representa o loop de aprendizado de uma rede neural. O papel do supervisor está descrito no bloco otimizador. Um dos algoritmos de treinamento supervisionado é o da retropropagação do erro (*back-propagation*).

$$SSE = \sum (y - \hat{y})^2 \quad (4.6)$$

O algoritmo de retropropagação do erro data da década de 1980 quando se popularizou na ciência para a resolução do problema de atualização dos pesos sinápticos para o perceptron de múltiplas camadas através do processo de treinamento supervisionado,

Fonte: adaptado de Chollet (2007)

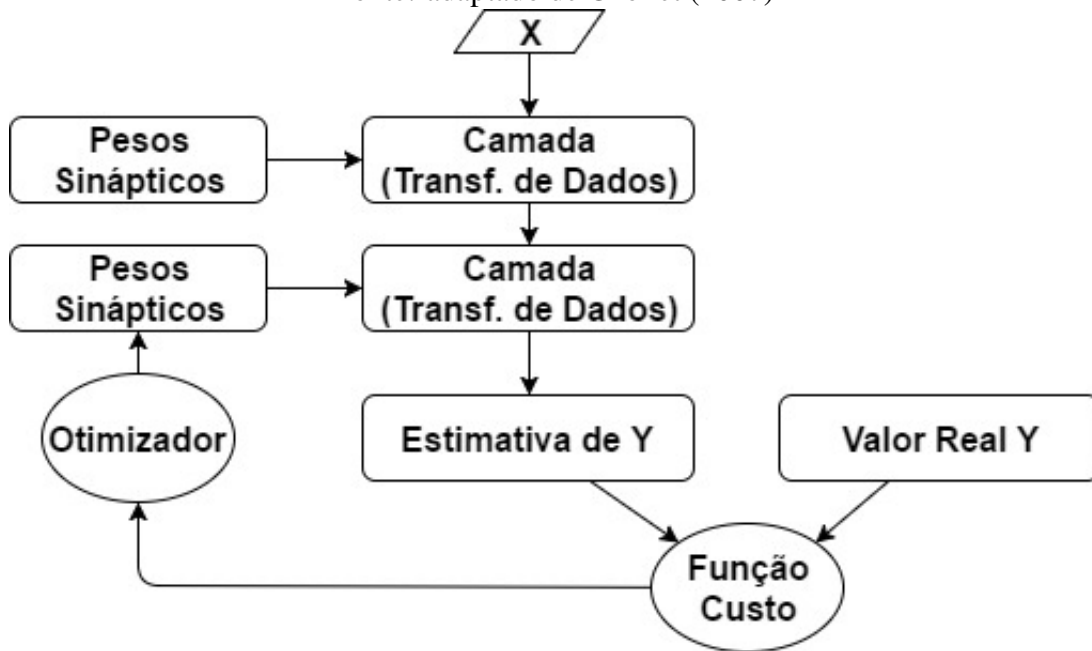


Figura 4.5: Processo de Atualização dos Pesos Sinápticos e aprendizado neural.

pois, sabendo que a arquitetura de uma rede perceptron de múltiplas camadas contém nós ocultos, o supervisor não tem o controle do sinal do erro nas camadas mais internas.

Ele é composto por dois passos: para frente e para trás. O primeiro passo consiste na inserção do sinal de entrada nas primeiras camadas e a propagação da informação, através dos pesos sinápticos até a saída da rede. Já o caminho inverso, consiste em determinar o valor da diferença entre a resposta desejada (d_j) e a resposta real (y_j) do neurônio j , na iteração n (eq. 4.7).

$$\rho = d_j(n) - y_j(n) \quad (4.7)$$

O processo de atualização dos pesos sinápticos começa a partir do sinal de erro de saída, descrita na eq. 4.7, que por sua vez gera outros sinais de erro para as camadas mais internas. Dessa forma, o supervisor conseguiria modificar o peso das interconexões entre neurônios internos. O algoritmo da retropropagação objetiva ajustar os pesos de todos os neurônios de tal forma que a função custo, referente a energia do erro de saída, descrita pela Eq. 4.8, seja mínima (2001).

$$\varepsilon(n) = \frac{1}{2} \sum \rho^2(n) \quad (4.8)$$

Capítulo 5

Metodologia

Nesta sessão será descrito todo o procedimento para a obtenção do objetivo deste trabalho, desde a organização dos testes reais no veículo *off-road*, à captura dos dados através de um datalogger automotivo. Esses dados serão utilizados com o pós processamento para a análise de identificação *off-line* via Matlab.

5.1 AIM EV0 4 Data logger

A captura dos dados embarcados no carro foi realizada através de um datalogger profissional. O modelo é EVO 4 da fabricante italiana AIM (AiM Tech Srl, 2017), sendo comumente utilizado por equipes de Formula SAE, Karts e Bajas. Ele possui extrema robustez a intemperes do ambiente, podendo ser inserido em qualquer outro tipo de veículo. Além disso, seu hardware é extramente poderoso pois conta com cinco canais de entrada para sensores analógicos (e.g. transdutores lineares para curso de suspensão) e digitais (sensor de velocidade linear) além de interfaces de hardware para conexões de redes automotivas tais como CAN e K-Line. Mais ainda, o aparelho ainda conta com um GPS interno de alta precisão e acelerômetros para os eixos x, y e z. A Figura 5.1 retrata, em vista frontal, o data logger usado neste trabalho.

Da mesma forma, o seu software embarcado é capaz de filtrar os sinais e gravá-los na memória interna com formas de onda limpas. As variáveis disponíveis para a leitura estão descritas na tabela 5.1. Dentre elas, foram selecionadas apenas: RPM, Velocidade linear, aceleração lateral, aceleração longitudinal e velocidade angular.

Mais ainda, o programa responsável por fazer a leitura dos dados gravados pelo Datalogger e sua configuração interna é o *Race Studio 2*. Dentro dele, encontra-se o *Race Studio 2 Analysis* que é uma ferramenta de análise gráfica e tem extrema importância para a manipulação dos dados. Neste último, os dados de treino foram importados, analisados visualmente através de gráficos, podendo ser exportados para diversos formatos, inclusive

Fonte: (AiM Tech Srl, 2017)



Figura 5.1: AIM EVO 4 Data logger.

.mat (Matlab). O software está disponível, gratuitamente, na página do desenvolvedor do Datalogger. (AiM Tech Srl, 2017)

Fonte: elaborada pelo autor



Figura 5.2: Fixação do data logger no Baja.

O sistema DAS deve permanecer em um local imune a interferências eletromagnéticas (e.g. motor) e mecânicas (e.g. peças móveis) as quais podem distorcer a captura dos sinais analisados. Mais ainda, acelerômetros devem ficar tão próximos quanto for possível do centro de gravidade do veículo a fim de evitar valores com *offset*. Dessa forma, o data logger está fixado, através de parafusos em baixo do banco do piloto, ao assoalho do carro, como mostra a Figura 5.2. Conectado ao sistema DAS, está um fio que vem do primário da bobina de ignição e vai até o canal de entrada "RPM/K-line" para a medição das revoluções do motor. Além desse, é necessário a conexão com os polos positivos e negativos da bateria (12V) para a alimentação do sistema. Os dados foram descarregados

para o computador através de um cabo USB com o data logger. Além disso, a taxa de gravação dos dados é de 10 milissegundos.

Tabela 5.1: Variáveis disponíveis no AIM EVO 4

Engine [rpm]	Rotação por minuto do motor
Vertical_acc [g]	Aceleração Vertical (Z), de acordo com o acelerômetro
Lateral_acc [g]	Aceleração Lateral (Y), de acordo com o acelerômetro
Longitudinal [g]	Aceleração Longitudinal (eixo X), de acordo com o acelerômetro
Datalogger_Tem [°C]	Temperatura Interna do Datalogger
Battery [V]	Tensão de alimentação do Datalogger
GPS_Speed [km/h]	Velocidade linear do veículo, de acordo com o GPS
GPS_Nsat [#]	Número de satélites pareados
GPS_LatACC [g]	Aceleração Lateral (Y), de acordo com o GPS
GPS_LonAcc [g]	Aceleração Longitudinal, de acordo com o GPS
GPS_Slope [deg]	Inclinação de afarda (eixo X)
GPS_Heading [deg]	Orientação em relação ao Norte Magnético
GPS_Gyro [deg/s]	Velocidade angular em curvas de acordo com o GPS
GPS_Altitude [m]	Altitude
GPS_PosAccuracy [m]	Precisão do sinal de GPS

5.2 Organização do Procedimento

A organização da proposta deste trabalho inicia-se a partir da discussão do melhor traçado de pista e da seleção dos pilotos de Baja. A seleção do perfil de pilotos dentre a equipe Car-Kará UFRN seguiu o padrão proposto em KATZOURAKIS *et al.* (2012), onde sugere-se a escolha de pilotos experientes e inexperientes para a formação do conjunto de dados. Ao todo, foram selecionados seis pilotos diferentes, de tal forma que alguns participaram de mais um teste, enquanto que outros estiveram presentes em apenas dois deles. A fim de termos o maior número de amostras dos pilotos, foram selecionados apenas quatro deles, sendo dois pilotos experientes (P1 e P3), um com nível intermediário (P2) e um novato (P4). A Tabela 5.2 resume a organização do experimento para seleção dos pilotos.

Além disso, para a formação da pista de teste, foram selecionados alguns espaços gramados nas dependências do campus universitário da UFRN. Nesses foram feitos sete traçados, em dias diferentes, que serão chamados de pista 1, pista 2, ..., pista 7. Todas elas têm diferentes tipos de obstáculos como areia fofa, troncos e rampas. Cada um dos pilotos teve a possibilidade de rodar algumas voltas em cada pista e, após o término delas, foi feito o descarregamento dos dados para o laptop. A partir da aquisição de dados, uma

Tabela 5.2: Distribuição dos pilotos nas pistas.

	P1	P2	P3	P4
Pista 1	X	X	X	
Pista 2	X	X	X	
Pista 3	X	X	X	
Pista 4	X	X	X	
Pista 5	X	X		X
Pista 6	X	X	X	X
Pista 7	X	X	X	X

rede neural artificial realizou a identificação dos pilotos como P1, P2, P3 e P4, de maneira *off-line*.

Capítulo 6

Estudo de Caso

Neste capítulo, serão descritos os procedimentos realizados para a classificação dos pilotos do baja. Esse procedimento pode ser dividido em dois estágios: organização e montagem do conjunto de dados; e otimização dos parâmetros de criação da Rede Neural Artificial. Antes dos comentários sobre a Rede Neural utilizada, serão expostos alguns gráficos que podem ser entendidos como um estudo de caso para melhor entendimento das variáveis coletadas nos testes. A análise estatística vem a tona com a estimativa das densidades de probabilidade de cada piloto bem como as correlações entre as variáveis analisadas.

6.1 Aquisição e organização do Conjunto de Dados

A aquisição dos dados no veículo Baja foi realizada através do data logger industrial AIM em conjunto com o software proprietário Race Studio 2. Esse software foi responsável pela exportação dos dados das voltas dos pilotos, em formato de espaço de trabalho para Matlab. Esse processo foi repetido de maneira contínua para todas as pistas de cada piloto. Dentre todas as variáveis possíveis de serem exportadas, vistas na Tabela 5.1, foram selecionadas cinco. Elas estão descritas na Tabela 6.1.

Com os dados exportados pelos programa *Race Analysis 2* para o formato .mat, foi possível a inclusão dos treinos individuais no espaço de trabalho do Matlab. Como citado anteriormente, foram selecionadas apenas as variáveis de rotação do motor, velocidade linear, aceleração lateral, aceleração longitudinal e velocidade angular.

Em seguida, os dados são importados para tratamento no Matlab. Já nesse *software*, os dados referentes as diferentes pistas de um mesmo piloto são concatenados, transformando em uma matriz de células, na qual cada coluna representaria as diversas voltas de uma única pista e cada célula dessa coluna seriam as variáveis descritas anteriormente. Os dados também são normalizados a fim de se ter média zero e variância unitária, através

Tabela 6.1: Variáveis Utilizadas

Variável	Descrição
Engine [RPM]	Rev. por minuto do motor
GPS_LatAcc.[g]	Aceleração lateral de acordo com o GPS
GPS_LonAcc.[g]	Aceleração longitudinal de acordo com o GPS
GPS_Speed [km/h]	Velocidade linear
GPS_Gyro [deg/s]	Velocidade angular

da Eq. 3.7.

Uma vez com a matriz de todas as pistas de um único piloto, pega-se cada célula (que representa as variáveis utilizadas) dessa matriz e divide-se seu tamanho em 3 ou 5 seções. Em cada parte seccionada, são computadas algumas estatísticas listadas abaixo de todas as variáveis presentes na Tabela 6.1, sendo "i" variando de 1 até 3 ou 5 divisões. O processo de divisão do conjunto de dados em seções é similar ao encontrado em Van Ly *et al.* (2013).

- Média (setor i);
- Variância (setor i);
- Valor máximo (setor i);
- Valor mínimo (setor i);

Fonte: elaborada pelo autor

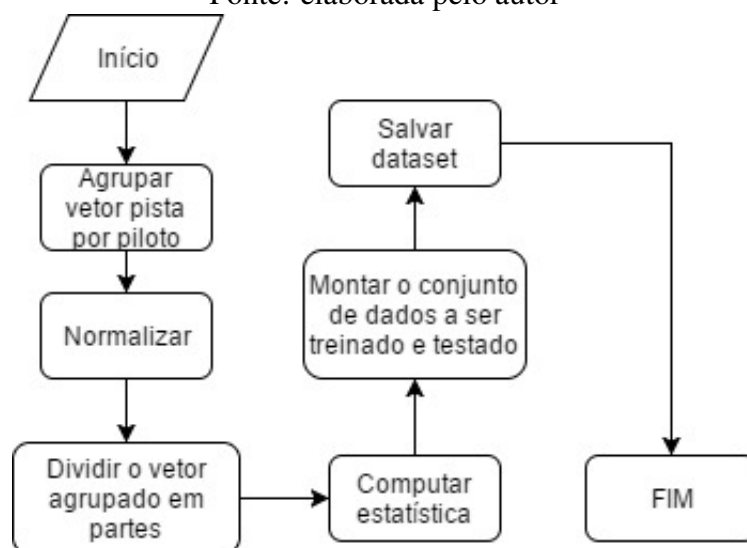


Figura 6.1: Fluxograma para a montagem do conjunto de dados.

Uma vez computada a estatística de cada setor, concatenam-se os vetores das estatísticas de tal forma que o seu agrupamento gera o vetor de características de cada piloto que será então entrada do modelo neural. A Figura 6.1 refere-se a sequência lógica para a criação do conjunto de dados de treinamento e teste utilizados na rede neural. O procedimento descrito é executado uma vez para cada piloto.

6.2 Análise Estatística dos Dados

Após a coleta dos dados, pode-se comentar sobre a sua respectiva análise estatística. Nesta seção serão expostos alguns tópicos sobre como é regido o sistema pilotos/pistas no que tange a estatística, sendo necessário o estudos através das estimativas das densidades de kernel de cada piloto em suas respectivas pistas, bem como a análise por correlação das variáveis presentes.

6.2.1 Estimativas das Densidades de Kernel

Como descrito na seção 5.2, serão analisados, a princípio, quatro pilotos em sete pistas distintas. A estimativa da distribuição de probabilidade pode ser encontrado na figura 6.2. Além delas, tem-se a Tabela 6.2 que destaca os valores médios e as variâncias das variáveis de cada piloto nas duas pistas de testes apresentadas neste trabalho.

Tabela 6.2: Média e Variância dos treinos dos 4 pilotos em 7 pistas.

	Média			
	Piloto 1	Piloto 2	Piloto 3	Piloto 4
Motor	2901.16	2869.61	2869.11	2791.7
Velo Linear	20.25	20.18	21.07	24.74
Acc. Lateral	0.128	0.079	0.095	0.01
Acc. Long.	0.005	0.006	0.006	0.001
Velo Angular	9.29	8.5	8.16	2.76
	Variância			
	Piloto 1	Piloto 2	Piloto 3	Piloto 4
Motor	130539.1	132480.7	115399.6	63639.57
Velo Linear	96.35	81.1	86.62	68.98
Acc. Lateral	0.26	0.149	0.222	0.112
Acc. Long.	0.02	0.021	0.019	0.017
Velo Angular	1910.13	1730.67	1812.29	928.57

Analisando a forma de distribuição da variável de revoluções do motor, percebe-se que os pilotos 1 e 2 tiveram comportamentos similares para valores maiores que 2800,

aproximadamente. No entanto, há uma pequena diferença no primeiro quartil. Além disso pode-se ver através da Tabela 6.2 que o primeiro piloto tem o maior valor médio encontrado, dentre os quatro indivíduos, para a variável em questão. Por outro lado, o piloto 4 tem a menor variância dos dados que justifica-se pelo menor número de amostras (apenas três pistas).

A análise estatística da velocidade linear é caracterizada por uma distribuição de probabilidade similar entre os pilotos 2 e 3 para valores entre 0 e 25km/h. Além disso, a maior parte dos valores de velocidade linear, quando comparados todos os pilotos, também se encontra nesse mesmo intervalo. Mais ainda, o maior valor de velocidade média foi encontrado no conjunto de dados do piloto 4 que é explicado por duas pistas rápidas que o mesmo participou (track 6 e 7) das três totais.

As distribuições de probabilidade para a aceleração lateral dos quatro pilotos são bem parecidas entre si. No entanto, os pilotos 1, 2 e 3 possuem valores médios próximos de 0,1g enquanto que o valor para o piloto 4 fica na casa dos 0,01g. Na análise por dispersão dos dados, os pilotos 1 e 3 possuem variância em torno de 0,240. Em contrapartida, os valores de variância para os pilotos 2 e 4 giram em torno de 0,120.

Além disso, analisando a aceleração longitudinal, tem-se semelhança para todos os pilotos. Os valor médio fica em torno de 0, enquanto que a variância está próxima do valor 0,02.

A Análise da variável velocidade angular chega-se a uma forma de distribuição de probabilidade em que os três primeiros pilotos possuem comportamento similares entre si. Os seus valores médios giram em torno de 8.16 a 9.29 graus/s. Mais ainda, os valores de variância dos dados ficam entre 1730 e 1910. Por outro lado, a forma de distribuição para o piloto 4 possui um certo grau de simetria, quando comparados com os outros três, tendendo a uma distribuição normal com valor médio igual a 2,76 graus/s e variância 928,57.

Em adição, realizou-se a análise de correção das variáveis com o intuito do melhor entendimento do comportamento do sistema piloto + carro.

6.2.2 Análise de Correlação dos Dados

A análise de correlação entre as variáveis com o conjunto de dados foi realizada envolvendo todos os pilotos nas 7 pistas. O estudo foi desenvolvido através da matriz de correlação, retratada pela Figura 6.3. Nela, podem-se perceber algumas principais características do comportamento de um veículo do tipo Baja SAE. A mais fácil de ser identificada é a tendência positiva entre as variáveis de velocidade angular e aceleração lateral.

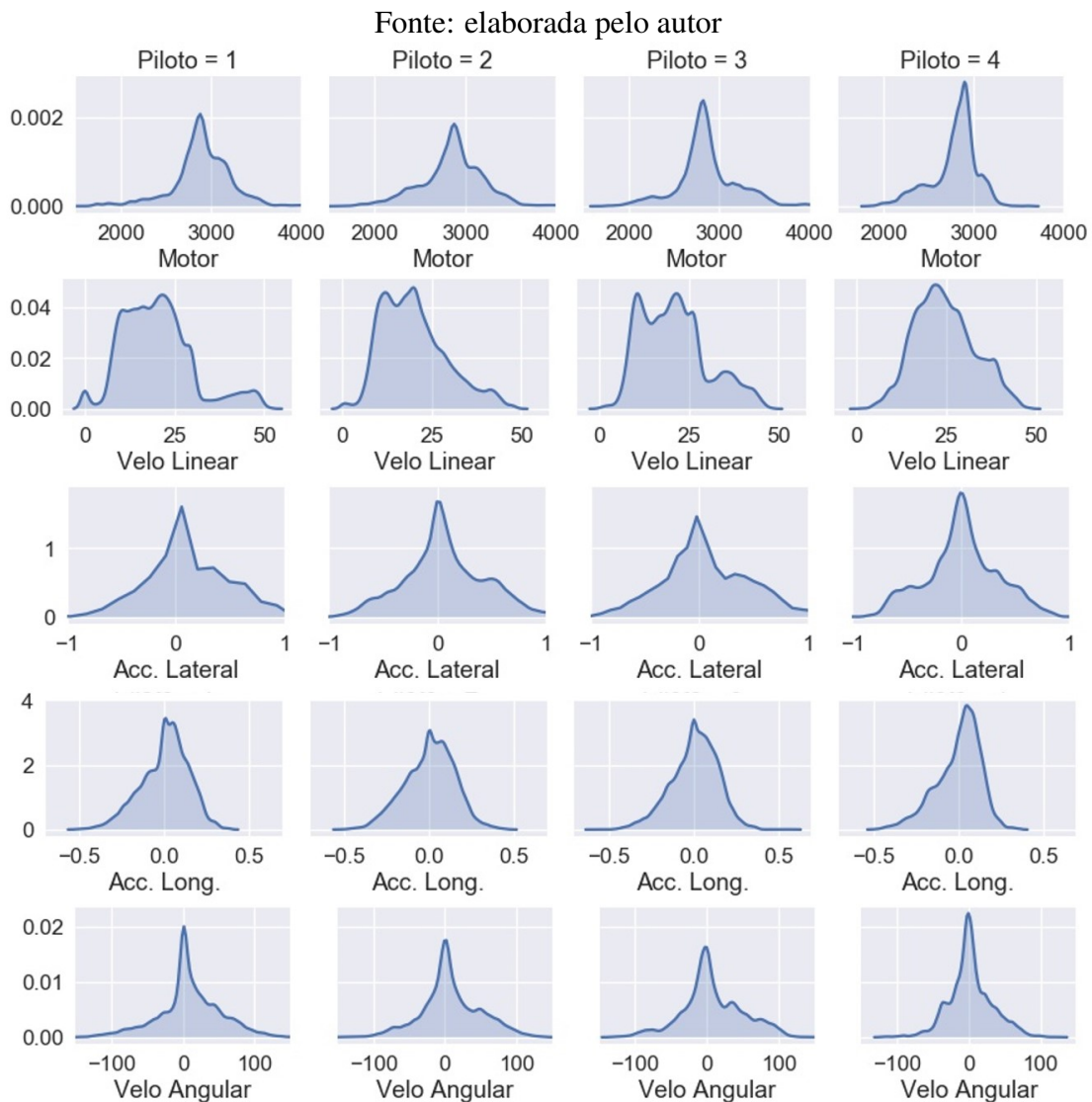


Figura 6.2: Distribuição Estimada de Kernel dos pilotos nas diversas pistas

Ou seja, quanto maior forem os valores de uma, maiores tenderão os da outra respectiva. Por outro lado, algumas correlações negativas ocorrem entre os pares de variáveis: velocidade linear e aceleração lateral; velocidade linear e velocidade angular; aceleração lateral e aceleração longitudinal. Os outros pares de variáveis possuem desconexão, em outras palavras, a tendência de crescimento ou redução de uma não afeta a tendência da outra.

Em seguida serão feitos alguns comentários sobre a criação da rede neural de múltiplas camadas para a classificação dos quatro pilotos.

Fonte: elaborada pelo autor

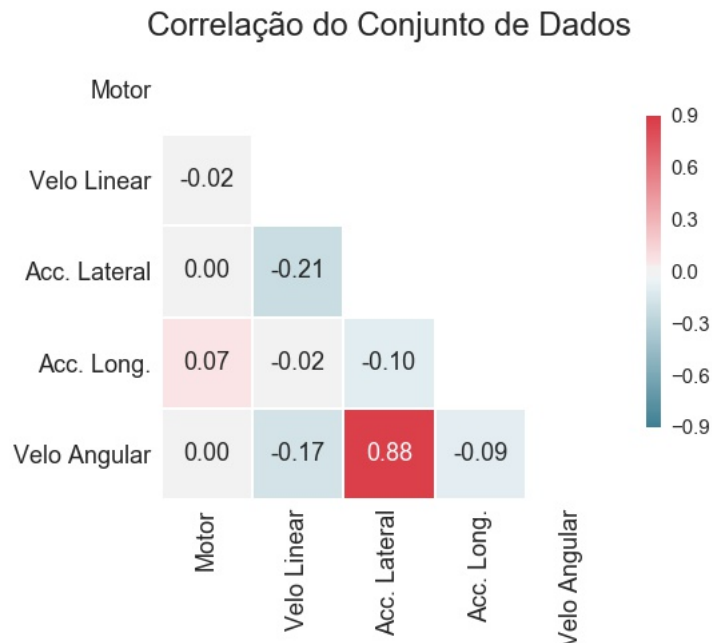


Figura 6.3: Correlação das variáveis para o conjunto de dados utilizando amostras de todos os pilotos nas 7 pistas.

6.3 Classificação dos Pilotos via Rede Neural

Após o conjunto de dados ser salvo, entra-se no segundo estágio. Aqui, faz-se a importação dos dados e divide-os entre dados para treinamento da rede neural (75%) e dados para teste (25% de amostras não vistas anteriormente).

A fim de se encontrar os melhores parâmetros da rede neural (número de neurônios nas camadas ocultas), foram criados dois *scripts*: um que incrementava apenas o número de neurônios da primeira camada oculta, chamando de H1, e outro que incrementava tanto primeira cada oculta (H1) quanto a segunda camada (H2).

Para o primeiro *script* a variável H1 foi iniciada com valor de 10 e seu valor final foi definido como 80, o passo incremental é de 2 neurônios. Já para o outro *script*, determinou-se o valor inicial de $H1_{inicial} = 10$, o valor final $H1_{final} = 80$ e passo incremental de 4. Da mesma maneira, determinado-se o valor inicial de H2 como 2 e valor final como 50, com passo incremental, aqui, de 4 unidades. Essa escolha foi tomada empiricamente.

A rede então é treinada através do algoritmo *Gradient Descent Backpropagation*,

utilizando-se dos valores atuais de números de neurônios ocultos H1 e H2. As funções de ativação utilizadas nas camadas ocultas são do tipo sigmóide, enquanto que a camada de saída utiliza-se da *softmax*, por ser um problema de classificação (BISHOP, 1995). A função custo utilizada é o *erro médio quadrático*. Além disso, é utilizada a parada antecipada, de atualização dos pesos sinápticos, por erro de validação, a fim de aprimorar a taxa de generalização do classificador.

Para o primeiro *script* a rede é testada com amostras não vistas anteriormente e, assim, o erro de teste é verificado e, caso o valor seja o mínimo até então, a rede é salva e o número H1 é incrementado, repetindo-se o ciclo. O fluxograma que representa o processo treinamento e teste da rede neural que otimiza apenas a primeira camada oculta está retratado na Figura 6.4.

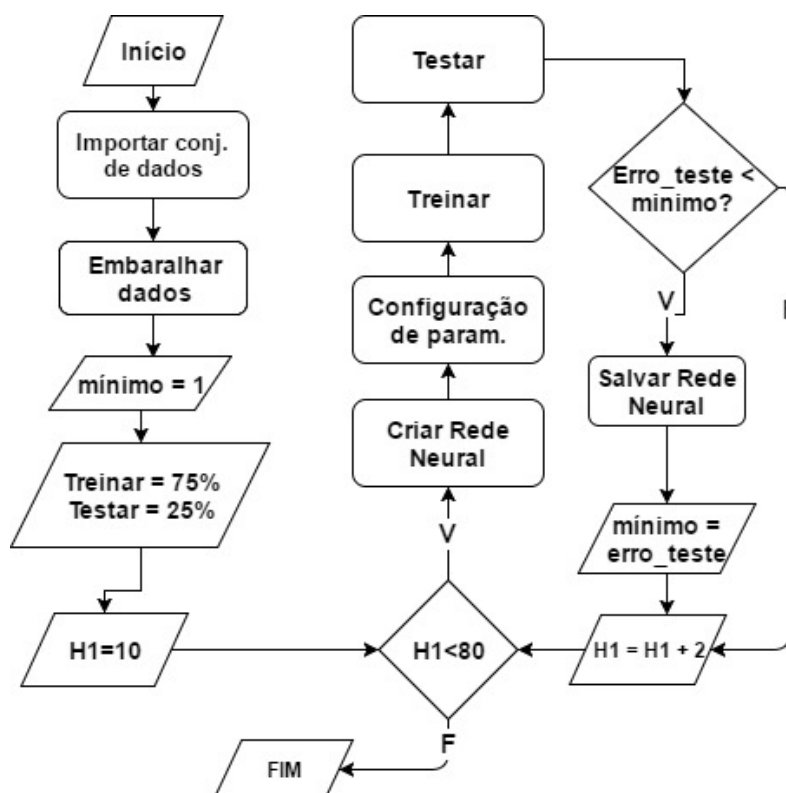


Figura 6.4: Fluxograma do processo de treinamento da primeira camada oculta da rede neural.

O mesmo ocorre quando executado o *script* que otimiza a segunda camada oculta neurônios. A rede é testada com amostras não vistas anteriormente e se o erro de teste for menor que o mínimo, a rede é salva e o valor de H2 é incrementado para a próxima iteração. Quando H2 atinge o valor máximo o valor de H1 passa a ser incrementado também e novamente o ciclo se repete. O fluxograma que representa o processo de treinamento e

teste da rede neural que otimiza apenas as camadas ocultas 1 e 2 está retratado na Figura 6.5.

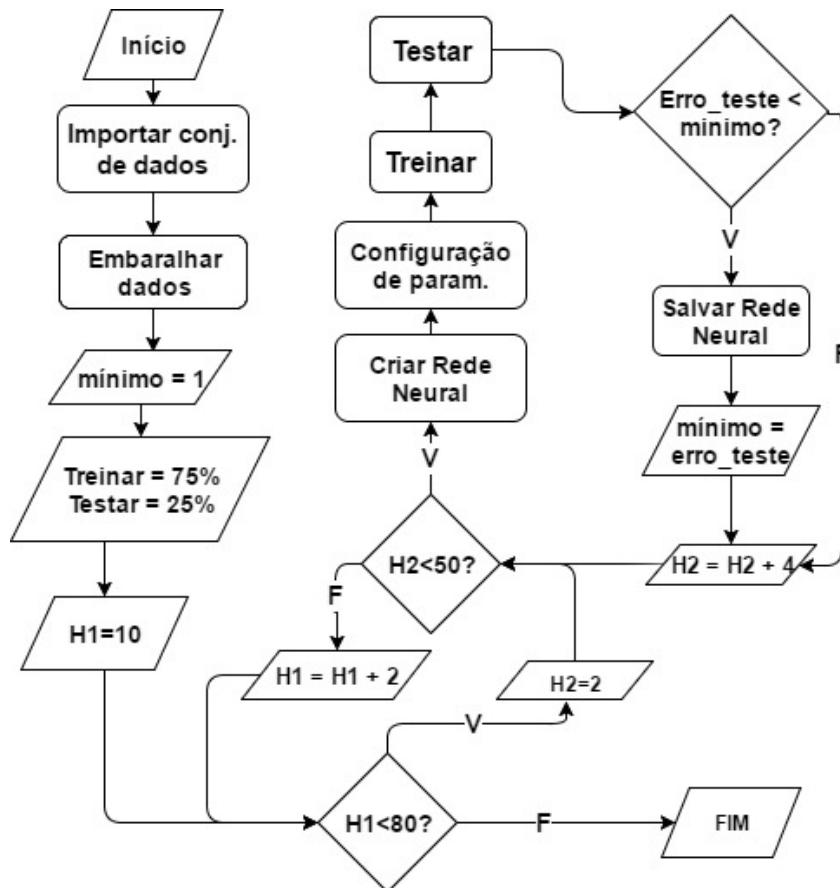


Figura 6.5: Fluxograma do processo de treinamento das camadas ocultas 1 e 2 da rede neural.

Uma vez executado o *script* de otimização, é chegado o momento de verificar a consistência dos parâmetros. Dessa forma, roda-se um outro programa que importa a rede neural salva e retesta-a com o conjunto de dados embaralhados. Em resumo, tem-se a Tabela 6.3, que descreve explicitamente as características envolvidas para a implementação da rede neural proposta neste trabalho.

Tabela 6.3: Resumo das características da rede neural implementada

Característica	Descrição
Entrada	Vetor das estatísticas dos setores (1,...,3,...5)
Número de Entrada	100 para 5 setores / 60 para 3 setores
Saída	Identificação de piloto como P1, P2, P3 e P4
Número de saídas	2 (Binária - 00,01,10,11)
H1	1,...,80 neurônios
H2	1,...,50 neurônios
Algoritmo de Treinamento	<i>Gradient Descent Backpropagation</i>
Validação	Parada antecipada por 6 testes

Capítulo 7

Resultados

Utilizando-se do procedimento descrito anteriormente, quatro experimentos foram testados: dois com a divisão das células do vetor de características para 3 e 5 partes através da rede neural treinada apenas com uma camada oculta; outros dois testes foram feitos com a rede neural com duas camadas ocultas.

Para o primeiro tipo de experimento, não se verificou uma taxa de classificação satisfatória. Encontrou-se que 87,5% das amostras foram classificadas corretamente para a divisão do vetor de características em 5 partes. Mais ainda, apenas 81% de amostras foram classificadas com êxito quando testado o conjunto de dados a partir da divisão em 3 partes do vetor de características.

No entanto, para o segundo tipo experimento, executou-se o *script* de busca dos melhores parâmetros e foi encontrado que para valores $H1=34$ e $H2=26$ tem-se uma taxa de sucesso em torno de 97%, para a divisão do vetor de características em 3 partes. Enquanto que quando executado o teste em que as variáveis do vetor de características são divididas em 5 partes, encontrou-se um novo conjunto ótimo de parâmetros, sendo $H1=64$ e $H2=22$. Com esse experimento, a taxa de sucesso para o conjunto de teste ficou em torno de 93%.

É válido ressaltar que a melhor a configuração para o teste em uma única de camada ocorreu quando o vetor de características foi dividido em 3 partes. Esse mesmo padrão foi encontrado para o *script* de otimização das duas camadas ocultas.

Mais ainda, a Figura 7.1 retrata as Densidades de Distribuição de Probabilidade dos erros de teste para os experimentos em 3 e 5 divisões da rede neural de múltiplas camadas. É possível perceber que para o teste com 3 divisões, a probabilidade de ocorrência dos erros entre 0 e 20% é maior do que quando comparado com o teste em 5 divisões.

Apesar de um reduzido tamanho de amostras, esse resultado demonstra que o metodologia utilizada mostrou ser efetiva quando comparada com os resultados encontrados por (Van Ly *et al.*, 2013), que utilizando o método k-means obteve uma taxa de sucesso em torno de 80% e 60% para o método de Máquinas de Vetor de Suporte (SVM).

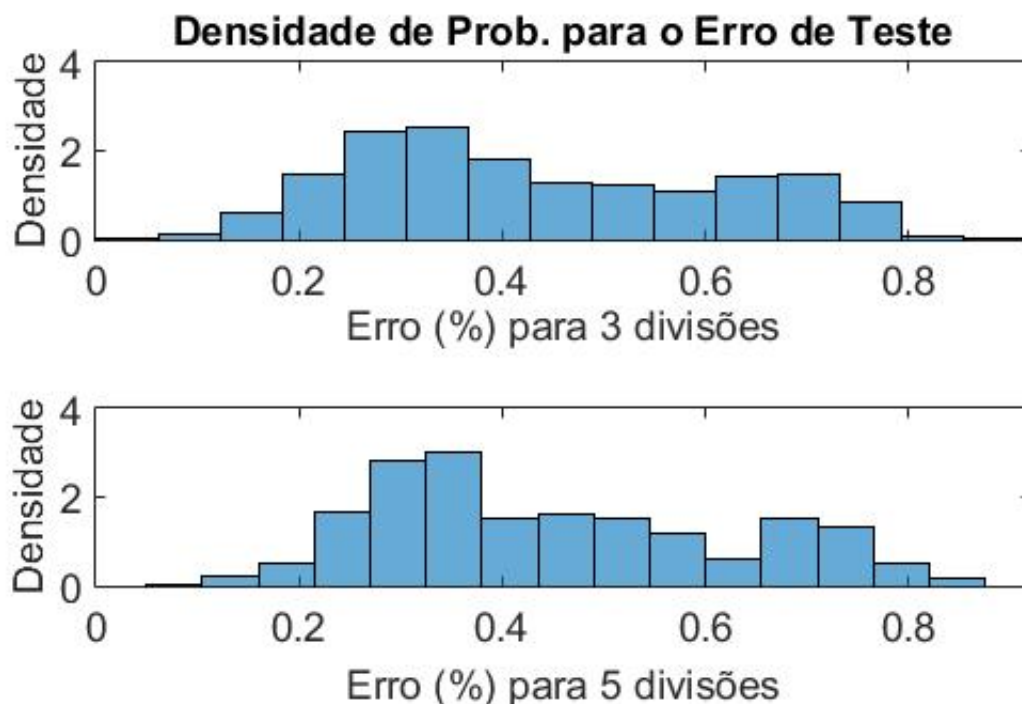


Figura 7.1: Densidades de probabilidade dos erros de teste

Além disso, foram executados os mesmos procedimentos acima com o algoritmo de treinamento *Levenberg-Marquardt*. Mediu-se o tempo de treinamento e fez-se uma comparação com o algoritmo *Gradient Descent*. Os valores estão descritos na Tabela 7.1. Nela, pode-se constatar que o algoritmo de *Levenberg-Marquardt* possuiu um tempo de treinamento elevado quando comparado com o algoritmo do *Gradient Descent* para todos os casos possíveis.

Tabela 7.1: Comparativo entre os tempos de treinamento para os algoritmos de Levenberg-Marquardt e Gradient Descent.

Num. Camadas Ocultas	Num. De Seções	Algoritmo	Tempo (s)
H1	3 Divisões	<i>Gradient Descent</i>	43.9
		<i>Levenberg-Marquardt</i>	344.2
	5 Divisões	<i>Gradient Descent</i>	57.2
		<i>Levenberg-Marquardt</i>	1.26E+03
H1+H2	3 Divisões	<i>Gradient Descent</i>	605.5
		<i>Levenberg-Marquardt</i>	1.04E+04
	5 Divisões	<i>Gradient Descent</i>	561.2
		<i>Levenberg-Marquardt</i>	2.24E+04

Por fim, ainda foram executados testes de classificação com combinações dos seis

pilotos tomadas duas a duas, ou seja, foram criadas combinações entre P1 x P2, P1 x P3, P1 x P4 e assim sucessivamente. Esse experimento é semelhante ao feito em QUEK e NG (2013). A taxa de acerto nas combinações demonstraram bastante satisfatórias tanto para a divisão em 3 quanto em 5 sessões, de tal forma que obteve-se 100% de classificação em ambos os casos, para a maioria das combinações entre pilotos. Apenas para o teste entre os pilotos 2 e 3 que a taxa de acerto ficou em torno de 94% para a divisão de 3 setores e em torno de 97% para a divisão de 5 setores. A Tabela 7.2 sumariza os resultados obtidos neste trabalho.

Tabela 7.2: Taxa de classificação da Rede Neural para o experimento dos quatro pilotos.

Nº de camadas ocultas	Divisões	Sucesso
Uma camada	5 divisões	81.30%
	3 divisões	87.50%
Duas camadas	5 divisões	93.80%
	3 divisões	96.90%

Capítulo 8

Conclusão

Em resumo, com o uso da metodologia proposta neste trabalho, uma alta taxa de sucesso na classificação, tanto nas combinações de pilotos tomados dois a dois, quanto na classificação geral utilizando o conjunto de dados dos quatro indivíduos, foi alcançada.

É válido ressaltar que o teste da rede neural com apenas uma única camada oculta demonstrou performance média inferior quando comparado ao teste executado em uma rede neural com duas camadas ocultas. Esse comportamento é esperado devido a melhor extração de características do modelo.

Os resultados alcançados a partir de duas camadas ocultas de neurônios e a divisão do vetor de características em 3 partes demonstraram um aumento em 20% de classificações quando comparado com outras propostas semelhantes a este trabalho (Van Ly *et al.*, 2013), mesmo utilizando-se de um cenário *off-road*. Dessa forma, confirmam-se que os testes foram satisfatórios e atingiu-se o objetivo proposto.

Alguns trabalhos futuros em relação ao *feedback* de comportamento de direção aos pilotos envolvidos são propostos. Eles podem ser entendidos como diminuição/aumento do ritmo de corrida, previsão de tempo de parada nos boxes, entre outros. Por exemplo, em um estágio mais simples podem ser realizados novos testes, com mais pilotos de diferentes níveis, para verificar a acurácia da rede quando a saída for o nível de experiência (veterano, intermediário ou iniciante). Isso pode ser utilizado em clima de competição para ajustar o tipo de performance do piloto e assim consumir mais ou menos combustível, dependendo do objetivo da equipe.

Além disso, propõe-se também o aprimoramento do sistema de aquisição de dados do veículo da Equipe Car-Kará, desenvolvido em NUNES (2016), a fim de confrontar as informações adquiridas pelo data logger da AIM Sports com a arquitetura de *hardware* desenvolvida até então. Caso as informações sejam consistentes, torna-se interessante rodar a rede neural a partir dos dados da eletrônica Car-Kará e comparar os resultados obtidos neste trabalho.

Como também é pertinente o teste com outros algoritmos de *Machine Learning*, encontrados na literatura, tais como: SVM, Regressão Logística, Camadas Ocultas de Markov, entre outros, e confrontar as performances.

Referências Bibliográficas

AGUIRRE, L. *Introdução à Identificação de Sistemas – Técnicas Lineares e Não-Lineares Aplicadas a Sistemas Reais*. Editora UFMG, 2015. ISBN 9788570415844. Disponível em: <<https://books.google.com.br/books?id=f9IwE7Ph0fYC>>.

AiM Tech Srl. *EVO 4 Datalogger*. 2017. Acessado em 27/09/2017. Disponível em: <<http://www.aim-sportline.com>>.

ANDRIA, G.; ATTIVISSIMO, F.; Di Nisio, A.; LANZOLLA, A. M.; PELLEGRINO, A. Development of an automotive data acquisition platform for analysis of driving behavior. *Measurement: Journal of the International Measurement Confederation*, Elsevier Ltd, v. 93, p. 278–287, 2016. ISSN 02632241. Disponível em: <<http://dx.doi.org/10.1016/j.measurement.2016.07.035>>.

Bill Goldwin. *Formula 1: Mercedes team turns to big data in quest for victory in Grand Prix*. 2017. Acessado em 07/02/2018. Disponível em: <<http://www.computerweekly.com/feature/Formula-1-Mercedes-team-turns-to-big-data-in-quest-for-victory-in-Grand-Prix/>>.

BISHOP, C. *Neural networks for pattern recognition*. [S.l.: s.n.], 1995. v. 103. 164–193 p. ISSN 0162-1459. ISBN 0198538642.

CHOLLET, F. *Deep Learning with Python*. [s.n.], 2007. v. 80. 453 p. ISSN 9781935182399. ISBN 9781937785536. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/20608803>>.

COELHO, A. A. R.; COELHO, L. dos S. *Identificação de sistemas dinâmicos lineares*. [S.l.: s.n.], 2004.

EREN, H.; MAKINIST, S.; AKIN, E.; YILMAZ, A. Estimating driving behavior by a smartphone. *IEEE Intelligent Vehicles Symposium, Proceedings*, n. June 2012, p. 234–239, 2012. ISSN 1931-0587.

FLEMING, W. J. Overview of automotive sensors. *IEEE Sensors Journal*, v. 1, n. 4, p. 296–308, Dec 2001. ISSN 1530-437X.

FLEMING, W. J. New automotive sensors - A review. *IEEE Sensors Journal*, v. 8, n. 11, p. 1900–1921, 2008. ISSN 1530437X.

Geoff Spencer. *Artificial Intelligence and Formula One: Bots on pole position in the race for technology*. 2017. Acessado em 07/02/2018. Disponível em: <<https://news.microsoft.com/en-my/2017/09/19/artificial-intelligence-formula-one-bots-pole-position-race-technology/>>.

GLANCY, D. J. Autonomous and automated and connected cars-oh my: first generation autonomous cars in the legal ecosystem. *Minn. JL Sci. & Tech.*, HeinOnline, v. 16, p. 619, 2015.

GRAMACKI, A. *Studies in Big Data 37 Nonparametric Kernel Density Estimation and Its Computational Aspects*. [S.l.: s.n.], 2018. ISBN 9783319716879.

HAYKIN, S. *Redes Neurais - Princípios e Prática [2ª Ed][Haykin][2001].pdf*. [S.l.: s.n.], 2001. 902 p. ISBN 85-7307-718-2.

KATZOURAKIS, D. I.; VELENIS, E.; ABBINK, D.; HAPPEE, R.; HOLWEG, E. Race-car instrumentation for driving behavior studies. *IEEE Transactions on Instrumentation and Measurement*, v. 61, n. 2, p. 462–474, 2012. ISSN 00189456.

LJUNG, L. *Ljung L System Identification Theory for User*. [s.n.], 1987. v. 25. 475–476 p. ISSN 00051098. ISBN 0138816409. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/0005109889900198>>.

MESEGUER, J. E.; CALAFATE, C. T.; CANO, J. C.; MANZONI, P. DrivingStyles: A smartphone application to assess driver behavior. *Proceedings - International Symposium on Computers and Communications*, p. 535–540, 2013. ISSN 15301346.

MESEGUER, J. E.; TOH, C. K.; CALAFATE, C. T.; CANO, J. C.; MANZONI, P. Drivingstyles: A mobile platform for driving styles and fuel consumption characterization. *Journal of Communications and Networks*, v. 19, n. 2, p. 162–168, 2017. ISSN 12292370.

MONTGOMERY, D. C.; PECK, E. A.; VINING, G. G. *Introduction to linear regression analysis*. [S.l.: s.n.], 1981. 504 p. ISBN 0470542810, 9780470542811.

NGO, V. D.; HOFMAN, T.; STEINBUCH, M.; SERRARENS, A. Gear shift map design methodology for automotive transmissions. *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, v. 228, n. 1, p. 50–72, 2014. ISSN 09544070.

NUNES, I.; SPATTI, D.; FLAUZINO, R. *Redes neurais Artificiais para Engenharia e Ciências Aplicadas*. [S.l.: s.n.], 2010.

NUNES, T. *Telemetria de um veículo Baja SAE através de rede CAN*. Natal: [s.n.], 2016.

PAPOULIS, A.; PILLAI, S. *Probability, random variables and stochastic processes*. [S.l.]: McGraw-Hill, 2002. (McGraw-Hill).

PISHRO-NIK, H. *Introduction to Probability, Statistics, and Random Processes*. [S.l.]: Kappa Research, LLC, 2014. ISBN 9780990637202.

QUEK, Z. F.; NG, E. Driver Identification by Driving Style. p. 1–4, 2013.

SAE Brasil. *Compeção Baja SAE Brasil - Regras e Relatórios*. 2018. Acessado em 16/12/2017. Disponível em: <<http://portal.saebrasil.org.br/programas-estudantis/baja-sae-brasil/regras>>.

SCHOREADER, L. D.; SJOQUIST, D. L.; E., S. P. *Understanding Regression Analysis An Introductory Guide Quantitative Applications in the Social Sciences*. Sage, 1986.

SILVERMAN, B. W. *Density Estimation for Statistics and Data Analysis*. [S.l.: s.n.], 1996. – p. ISBN 0412246201.

TAN, P.-N.; STEINBACH, M.; Vipin Kumar. *Introduction to data mining*. [S.l.: s.n.], 2006. 796 p. ISSN 00224405. ISBN 9789332518650.

TURLACH, B. *Bandwidth selection in kernel density estimation: A review*. [s.n.], 1993. 1–33 p. Disponível em: <<http://www.stat.washington.edu/~ebfox/courses/stat527/s13/readings/Turlach.p>>.

Van Ly, M.; MARTIN, S.; TRIVEDI, M. M. Driver classification and driving style recognition using inertial sensors. *IEEE Intelligent Vehicles Symposium, Proceedings*, n. Iv, p. 1040–1045, 2013. ISSN 1931-0587.

ZHANG, D. *Vehicle Parameters Estimation and Driver Behavior Classification for Adaptive Shift Strategy of Heavy Duty Vehicles*. 2017.

ZHANG, Y.; LIN, W. C.; CHIN, Y. K. S. A pattern-recognition approach for driving skill characterization. *IEEE Transactions on Intelligent Transportation Systems*, v. 11, n. 4, p. 905–916, 2010. ISSN 15249050.